OPTIMUM ESTIMATION OF A DELAY VECTOR CAUSED BY A RANDOM FIELD PROPAGATING ACROSS AN ARRAY OF NOISY SENSORS

William R. Hahn

Naval Ordnance Laboratory
White Oak, Maryland

16 June 1972

NOLTR 72-120

OPTIMUM ESTIMATION OF A DELAY VECTOR CAUSED BY A RANDOM FIELD PROPAGATING ACROSS AN ARRAY OF NOISY SENSORS

By
W. R. Hahn

16 JUNE 1972

NAVAL ORDNANCE LABORATORY, WHITE OAK, SILVER SPRING, MARYLAND

NOLTR 72-120

UNCLASSIFIED

U.S. NAVAL ORDNANCE LABORATORY

WHITE OAK

SILVER SPRING, MARYLAND 20910

AD 751555

To all holders of TR 72-120

Change 1

Title: OPTIMUM ESTIMATION OF A DELAY VECTOR CAUSED BY A RANDOM FIELD PROPAGATING ACROSS AN ARRAY OF NOISY SENSORS

1 Dec 1972

Approved by Commander, U.S. NOL

1 page(s)

By direction

This publication is changed as follows:

1. Change equation (2.1) to read:
$$X_i(\omega_k) = \frac{1}{\sqrt{T}} \int_T x_i(t) \exp\{-jk\omega_o t\}dt.$$

2. In equation (2.9) change $\delta_{mn} T$ to δ_{mn}.

3. In equation (2.10) change $\frac{T^2}{2\pi}$ to $\frac{T}{2\pi}$.

4. Change equation (2.42) to read
$$D \equiv (D_2, D_3, \ldots, D_M)^T.$$

5. In the first two lines of equation (3.20) change $\sum$ to $\sum_i$ once on each line.

6. In the first line of equation (3.22) change $2\ {}^2G^2$ to $2\omega^2G^2$.

7. In equations (4.10) and (4.11) change $\frac{1}{T}$ to $\frac{1}{\sqrt{T}}$.

8. In equation (4.12) change $\frac{1}{T^3}$ to $\frac{1}{T^2}$ and change $\frac{1}{T^2}$ to $\frac{1}{T}$.

9. In equation (4.14) change $\frac{1}{T^2}$ to $\frac{1}{T}$.

10. In equation (4.15) change $\frac{1}{T^4}$ to $\frac{1}{T^2}$.

11. In equation (4.18) delete T in two places.

12. In equation (A 14) change 0 to 1.

13. In line 5 of page 65 change $\begin{matrix} A & -B \\ B & A \end{matrix}$ to $\begin{bmatrix} A & -B \\ B & A \end{bmatrix}$.

Insert this change sheet between the cover and the title page of your copy.
Write on the cover "Change 1 inserted"

UNCLASSIFIED

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Naval Ordnance Laboratory White Oak, Silver Spring, Maryland 20910	UNCLASSIFIED
	2b. GROUP

3. REPORT TITLE

Optimum Estimation of a Delay Vector Caused by a Random Field Propagating Across an Array of Noisy Sensors

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

William R. Hahn

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
16 June 1972	76	13

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO.	NOLTR 72-120
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY

13. ABSTRACT

The problem of optimally processing multiple sensor data to determine the set of time delays generated by the propagation across an array of the wave fronts from a distant wide-band Gaussian noise source is investigated. It is assumed that the amplitude gradient across the array of the noise field is negligible, that the array outputs are corrupted by additive wide-band Gaussian independent sensor noises, and that the observation time is long. The Fisher Information Matrix is determined, and then used to show that the maximum likelihood estimate is asymptotically efficient (as theory dictates it should be). It is also shown that filtered correlator systems can provide asymptotically efficient estimates. Finally, the effects of suboptimal filtering of the inputs to a correlator system are investigated for the case when the signal and additive noise spectra are all band limited and have constant slopes of 0, -3, or -6 dB/octave.

DD FORM 1 NOV 68 1473 (PAGE 1)

S/N 0101-807-6801

UNCLASSIFIED
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Passive Multidimensional Sonar Localization						
Passive Sonar Signal Processing						
Optimum Filtering and Estimation						

NOLTR 72-120

OPTIMUM ESTIMATION OF A DELAY VECTOR CAUSED BY A RANDOM FIELD PROPAGATING ACROSS AN ARRAY OF NOISY SENSORS

Prepared by:
W. R. Hahn

ABSTRACT: The problem of optimally processing multiple sensor data to determine the set of time delays generated by the propagation across an array of the wave fronts from a distant wide-band Gaussian noise source is investigated. It is assumed that the amplitude gradient across the array of the noise field is negligible, that the array outputs are corrupted by additive wide-band Gaussian independent sensor noises, and that the observation time is long. The Fisher Information Matrix is determined, and then used to show that the maximum likelihood estimate is asymptotically efficient (as theory dictates it should be). It is also shown that filtered correlator systems can provide asymptotically efficient estimates. Finally, the effects of suboptimal filtering of the inputs to a correlator system are investigated for the case when the signal and additive noise spectra are all band limited and have constant slopes of 0, -3, or -6 dB/octave.

U. S. NAVAL ORDNANCE LABORATORY
WHITE OAK, MARYLAND

NOLTR 72-120 16 June 1972

OPTIMUM ESTIMATION OF A DELAY VECTOR CAUSED BY A RANDOM FIELD PROPAGATING ACROSS AN ARRAY OF NOISY SENSORS

This report is concerned with the theoretical bounds on the accuracy of estimation of the set of time delays caused by the propagation across a noisy array of the signal wave fronts from a distant random noise source. The Fisher Information Matrix is determined and used as a bench mark relative to which the efficacies of the maximum likelihood processor and a multiple multiplier correlator system are compared. The report will be of interest to those concerned with passive sonar detection and localization or the corresponding problem in seismology.

In its original form, this paper was submitted to the Graduate Faculty of the University of Maryland in partial fulfillment of the requirements for the degree of Doctor of Philosphy in Electrical Engineering. This graduate work at the University of Maryland received no NOL support.

The research reported herein is related to work done at the Naval Ordnance Laboratory, under Naval Ship Systems Command Task Number 08692/SF11-121-101, and is published as an NOLTR for information purposes.

ROBERT WILLIAMSON II
Captain, USN
Commander

E. H. BEACH
By direction

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

In many physical problems of interest, with sonar, radar, and seismology as examples, the time records of the outputs of an array of sensors are observed over some time interval to be used to estimate the position of a distant noise source. Typically, the outputs (i.e., time records) of the sensors consist of amplitude scaled and delayed replicas of the waveform from the distant source, corrupted by additive noises, usually local in origin. When the amplitude gradient of the noise from the distant source is negligible over the array, essentially all of the geometric information, e.g., range and bearing, is encoded in the set of time delays associated with the propagation across the array of the signal wavefronts from the distant source. This thesis treats the topics of filtering and signal processing to optimize the estimation of the time delays.

The distant noise source and the additive corrupting sensor noises are all assumed to be independent stationary Gaussian random processes. The time records to be processed are long compared to the signal and noise correlation times and also to the time needed for a signal wavefront to propagate across the array. These assumptions make an analysis based on a Fourier representation of the time records particularly convenient.

The Cramér-Rao matrix bound (CRMB) for the vector delay estimate is developed in Chapter II, and it is used as a benchmark, relative to which the efficacies of two estimation schemes are measured. These schemes are the maximum likelihood precessor in Chapter III, and a possibly suboptimal multiple correlator processor in Chapter IV. Theory dictates that the maximum likelihood processor should achieve the matrix bound, and this is shown. If the number of sensors is two (2) or three (3), then it is shown that the multiple correlator processor with optimally filtered inputs also achieves the matrix bound. The optimal filters are discussed in Chapter IV. Numerical results showing the effects of suboptimal filtering are presented in Chapter V for the specific case where the signal and noise spectra are band-limited with spectral slopes of 0, -3, or -6 dB/octave.

Although there have been many applications of the Cramér-Rao estimation theory to communication theory, they have generally been to problems of the known signal in Gaussian noise type. Among the investigators who have applied the Cramér-Rao estimation theory to problems in which the signals and noises all are random are Levin [1], Harger [2], and MacDonald and Schultheiss [3]. Levin investigated random process power spectrum parameter estimation. Harger treated the problem of optimally processing data of unknown focus in electro-optical systems. MacDonald and Schultheiss considered the problem of optimally estimating the bearing

fr m a sonar array to a distant noise source. In [3] the authors assume that the signal wavefronts at the array are planar, and the problem is essentially that of estimating a scalar. This thesis generalizes the problem treated by MacDonald and Schultheiss to the extent that the signal wavefronts across the array are assumed to have a curvature, and it is desired to determine this curvature, and hence the rest of the geometry, by measuring the vector of time delays associated with the propagation of the curved wavefronts across the array of sensors. This passive multidimensional localization problem appears not yet to have been treated in the literature.

The following notation is used in this thesis. If A is a matrix, then A^{-1} is its inverse, A* is its complex conjugate, A^T is its transpose, tr A is its trace, and det A is its determinant. If matrices A and B are positive definite, then $A > B (A \geq B)$ denotes that the difference A-B is positive definite (nonnegative definite). A square matrix whose elements off the main diagonal are all zero may be written as $\text{diag}(a_1, a_2, \ldots, a_n)$, where a_i is the i-th diagonal element. Vectors are column vectors unless otherwise specified. $\underline{1}$ denotes a vector with every element a one (1). $\underline{0}$ is a matrix of zeros, and I is the identity matrix. $\langle \cdot \rangle$ denotes the expectation operator, and grad f is the gradient of the scalar f. The gradient of a vector is the matrix in which the i-th row is the gradient of the

i-th component of the vector. The Kronecker delta is denoted in the usual fashion as δ_{ij}. If z is a complex number its real and imaginary parts are denoted by Re z and Im z, respectively.

The following conventions regarding integration limits are observed in this thesis. Integrals of the form $\int_{-\infty}^{\infty} f\,dx$ are written as $\int f\,dx$. Integrals of the form $\int_{-T/2}^{T/2} f\,dt$ are written as $\int_T f\,dt$. The quantity ω_N is defined in Chapter II. Integrals of the form $\int_{-\omega_N}^{\omega_N} f\,d\omega$ are written as $\int_B f\,d\omega$.

The symbols MLE, FIM, CRMB, and HOT are abbreviations for Maximum Likelihood Estimate, Fisher Information Matrix, Cramér-Rao Matrix Bound, and Higer Order Terms (as in series expansions), respectively. Appendix A is a brief introduction to the significance of the FIM and the CRMB.

Appendix B is an introduction to complex Gaussian random vectors, in the particular sense of Goodmann [4]. These complex Gaussian random vectors are pertinent to the study of the properties of real stationary Gaussian vector processes with long observation times.

CHAPTER II

THE LIKELIHOOD FUNCTION AND THE FISHER INFORMATION MATRIX

A. BASIC ASSUMPTIONS

In this chapter the likelihood function, $L(D)$, and the Fisher Information Matrix (FIM) are derived. The vector argument, D, of the likelihood function is the vector of delays to be measured. The following Assumptions are made:

1. The random signal and each of the additive sensor noises are all stationary zero-mean Gaussian random processes.

2. The stationary zero-mean Gaussian random processes of Assumption 1 are all independent.

3. The observation interval, T, is large compared to the correlation times of all the stationary zero-mean Gaussian random processes. It is also large compared to the time needed for the signal wavefronts to transverse the array.

B. THE LIKELIHOOD FUNCTION, L(D)

Each of the M sensors of the M-element array is observed for a duration of T seconds over the time interval $-\frac{T}{2} \leq t \leq +\frac{T}{2}$. The output of the i-th sensor is $x_i(t)$. The i-th time-record is represented by Fourier coefficients,

$$X_i(\omega_k) = \int_T x_i(t) \exp\{-jk\omega_o t\}dt, \qquad (2.1)$$

where $\omega_o = 2\pi/T$, and $\omega_k = k\omega_o$. If only frequencies up to $N\omega_o$ are to be processed, the set of M time records can be represented by MN complex Fourier coefficients. These MN coefficients can be treated as a single entity by defining a data vector X:

$$X^T = (X_1(\omega_1), X_2(\omega_1), \ldots, X_M(\omega_1), X_1(\omega_2), \ldots, X_M(\omega_N)) \qquad (2.2)$$

By Assumption 1, each of the components of the data vector is a zero-mean complex Gaussian random variable. Let the covariance matrix, R, be defined by

$$R \equiv < X^* X^T >. \qquad (2.3)$$

The elements of R are given by

$$X_p^*(\omega_m)X_q(\omega_n) = \int_T dt \int_T du \exp\{j(\omega_m t - \omega_n u)\} < x_p(t)x_q(u) >. \qquad (2.4)$$

It is assumed that the time record of the output of the p-th sensor is

$$x_p(t) = s(t-D_p) + n_p(t), \qquad (2.5)$$

where s(t) is the random signal waveform, D_p is the signal delay at the p-th sensor relative to the delay at an arbitrary reference point, and $n_p(t)$ is the independent additive sensor noise. Let $R_s(\tau)$ and $R_p(\tau)$ denote the signal and noise auto correlation functions, respectively. By Assumption 2 the expectation on the right side of Equation 2.4 is

$$< x_p(t)x_q(u) > = R_s(t-u+D_q-D_p) + \delta_{pq}R_p(t-u). \quad (2.6)$$

The autocorrelations are related to the spectral density functions $S(\omega)$ and $N_p(\omega)$ by

$$R_s(\tau) = \frac{1}{2\pi} \int S(\omega) \exp\{j\omega\tau\}d\omega \quad (2.7)$$

and

$$R_p(\tau) = \frac{1}{2\pi} \int N_p(\omega) \exp\{j\omega\tau\}d\omega \ . \quad (2.8)$$

Inserting Equations 2.6, 2.7, and 2.8 into Equation 2.4 and integrating with respect to t and u,

$$< X_p^*(\omega_m)X_q(\omega_n) > = \frac{T^2}{2\pi} \int d\omega \ (S(\omega)+N_p(\omega)) \exp\{j\omega(D_q-D_p)\} \cdot \left[\frac{\sin(\omega+2\pi m/T)T/2}{(\omega+2\pi m/T)T/2} \cdot \frac{\sin(\omega+2\pi n/T)T/2}{(\omega+2\pi n/T)T/2}\right]. \quad (2.9)$$

By assumption 3, the signal and noise correlation times as well as D_q-D_p are all small compared to T. The sinc functions [5] in Equation 2.9 are orthogonal, and since all the functions of ω inside the integral in Equation 2.9 vary slowly compared to the sinc functions, for practical purposes

$$\langle X_p^*(\omega_m) X_q(\omega_r) \rangle = \delta_{mn} T(S(\omega_n)+\delta_{pq}N_p(\omega_n)) \exp\{j\omega_n(D_q-D_p)\}. \tag{2.10}$$

The real and imaginary parts of X, say X_{RE} and X_{IM}, respectively, are each zero-mean Gaussian random vectors. If $\langle X_{RE}X_{RE}^T \rangle = \langle X_{IM}X_{IM}^T \rangle$ and if $\langle X_{RE}X_{IM}^T \rangle = -\langle X_{IM}X_{RE}^T \rangle$, then X is a complex Gaussian random vector, in the sense of Appendix B. This is true (asymptotically in T). Consider

$$\begin{aligned}
&\langle \mathrm{Re}\, X_p(\omega_m)\, \mathrm{Re}\, X_q(\omega_n) \rangle - \langle \mathrm{Im}\, X_p(\omega_m)\, \mathrm{Im}\, X_q(\omega_n) \rangle \\
&= \int_T dt \int_T du \langle x_p(t)x_q(u) \rangle \cos(\omega_m t+\omega_n t) \\
&= \frac{1}{2\pi} \int d\omega\, (S(\omega)+N_p(\omega)\delta_{pq}) \exp\{j\omega(D_p-D_q)\} \\
&\cdot \int_T dt \int_T du \exp\{j\omega(t-u)\} \cos(\omega_m t+\omega_n u) \\
&= \frac{1}{2\pi} \int d\omega\, (S(\omega)+N_p(\omega)\delta_{pq}) \exp\{j\omega(D_p-D_q)\} \\
&\cdot \frac{T^2}{2} \left[\frac{\sin(\omega+\omega_m)T/2}{(\omega+\omega_m)T/2} \; \frac{\sin(\omega-\omega_n)T/2}{(\omega-\omega_n)T/2} \right. \\
&\qquad \left. + \frac{\sin(\omega-\omega_m)T/2}{(\omega-\omega_m)T/2} \; \frac{\sin(\omega+\omega_n)T/2}{(\omega+\omega_n)T/2} \right] .
\end{aligned} \tag{2.11}$$

Again the sinc functions on the right side of the last part of Equation 2.11 are orthogonal. Since the integers m and n are positive only, the right side of Equation 2.11 is essentially zero. Thus the real and imaginary parts of the data vector X have the same covariance matrix. Almost identical arguments show that $\langle X_{IM}X_{RE}^T \rangle = -\langle X_{RE}X_{IM}^T \rangle$. It follows that the vector X is a complex Gaussian random vector (see Appendix B) and the density function for the vector is

$$p(X) = (\pi^{MN} \det R)^{-1} \exp\{X^T R^{-1} X^*\} . \qquad (2.12)$$

Through the remainder of this thesis the asymptotic nature of the results derived, particularly Equation 2.12, will be assumed, and only occasionally emphasized when it seems appropriate. Equations such as 2.12 will appear without apology.

Equation 2.10 shows that the Fourier coefficients for different frequencies are uncorrelated. It is therefore convenient to define the following vectors and matrices:

$$X(k) = (X_1(\omega_k), X_2(\omega_k), \ldots, X_M(\omega_k))^T \qquad (2.13)$$

$$V(k) = (e^{j\omega_k D_1}, e^{j\omega_k D_2}, \ldots, e^{j\omega_k D_M})^T \qquad (2.14)$$

$$N(k) = \mathrm{diag}(N_1(\omega_k), N_2(\omega_k), \ldots, N_M(\omega_k)) \qquad (2.15)$$

$$R(k) = N(k) + S(\omega_k)\, V^*(k)\, V^T(k) \qquad (2.16)$$

Using Equations 2.13 through 2.16 the density function can be written as

$$p(X) = \prod_{k=1}^{N} (\pi^M \det R(k))^{-1} \exp\{-X^T(k) R^{-1}(k) X^*(k)\} \qquad (2.17)$$

$$= [\pi^{MN} \prod_{k=1}^{N} \det R(k)]^{-1} \exp\{-\sum_{k=1}^{N} X^T(k) R^{-1}(k) X^*(k)\}.$$

An application of the "matrix inversion lemma" [6, p. 13] shows that

$$R^{-1}(k) = N^{-1}(k) - G(k) N^{-1}(k) V^*(k) V^T(k) N^{-1}(k), \qquad (2.18)$$

where G(k) is defined by

$$G(k) \equiv [1 + \sum_{k=1}^{M} S(\omega_k) N_i^{-1}(\omega_k)]^{-1} . \qquad (2.19)$$

The vector V(k) is conventionally called the steering vector.

In the remainder of the text it will be convenient to use the following simplified notation: The function G(k) may be written as G(k), or $G(k\omega_o)$, or $G(\omega_k)$, or $G(\omega)$ if $\omega = k\omega_o$ is understood, or simply as G with the argument suppressed if the argument is known. This notation will be used for all of the frequency dependent scalars, vectors, and matrices. In addition the set

$$B_+ = (i | i \text{ an integer, } 1 \leq i \leq N)$$

is defined in order that summations over the range of the Fourier frequencies, $\sum_{i=1}^{N}$, can be simply written as $\sum_{B+}$, whether the frequency arguments are or are not suppressed. This allows summations of the form $\sum_{i=1}^{M}$ to be written as $\sum_i$ and always understood to be an array sum.

The signal delay vector D is defined by

$$D = (D_1, D_2, \ldots, D_M)^T . \tag{2.20}$$

The vector D is the vector argument of the likelihood function, L(D). From Equation 2.17,

$$L(D) = [\pi^{MN} \prod_{B+} \det R]^{-1} \exp\{- \sum_{B+} X^T R^{-1} X^*\} \tag{2.21}$$

It is convenient at this point to examine det R(k), and, in particular, to show that it is not a function of D. Let $u = \exp\{j\omega_k\}$. Then

$$R(k) = \begin{bmatrix} S + N_1 & Su^{(D_2-D_1)} & \cdots & Su^{(D_M-D_1)} \\ Su^{(D_1-D_2)} & S + N_2 & \cdots & Su^{(D_M-D_2)} \\ \vdots & \vdots & & \vdots \\ Su^{(D_1-D_M)} & Su^{(D_2-D_M)} & \cdots & S + N_M \end{bmatrix} \quad (2.22)$$

From Equation 2.22, every element of the i-th column of R(k) has the factor u^{D_i}, and every element of the i-th row has the factor u^{-D_i}. Thus if R(k|i) is the R(k) matrix with D_i set to zero,

$$\det R(k) = (u^{-D_i})^M (u^{D_i})^M \det R(k|i). \quad (2.23)$$

In view of Equation 2.23 and the definition of R(k|i), det R(k) is not a function of D_i. Further,

$$\det R(k) = \det(N(k) + S(k)\underline{1}\,\underline{1}^T). \quad (2.24)$$

C. THE FISHER INFORMATION MATRIX

The CRMB for unbiased estimators of the vector argument Y is the inverse of the FIM, denoted by (FIM), where

$$(\text{FIM}) = - < \text{grad}(\text{grad } \ln L(Y))^T > . \qquad (2.25)$$

In Equation 2.25 L(Y) is the likelihood function for Y, and the gradients are taken with respect to the components of Y. The matrix $\text{grad}(\text{grad } \ln L(Y))^T$ is a matrix of second partial derivatives. From Equations 2.18, 2.19, 2.21, and 2.23,

$$\text{grad}(\text{grad } L(Y))^T = - \sum_{B+} \text{grad}(\text{grad } X^T R^{-1} X^*)^T$$

$$= \sum_{B+} G \, \text{grad}(\text{grad } X^T N^{-1} V^* V^T N^{-1} X^*)^T . \qquad (2.26)$$

Let a and b denote arbitrary elements of Y. The corresponding element of the FIM is determined by

$$- < \frac{\partial}{\partial a} \frac{\partial}{\partial b} \ln L(Y) > = - \sum_{B+} G < X^T N^{-1} \frac{\partial}{\partial a} (\frac{\partial}{\partial b} V^* V^T) N^{-1} X^* > . \qquad (2.27)$$

The k, m-th element of the matrix $V^* V^T$ is $\exp\{j\omega(D_m - D_k)\}$. Let

$$V^* V^T = (\exp\{j\omega(D_m - D_k)\}), \qquad (2.28)$$

so that

$$\frac{\partial}{\partial a}\left(\frac{\partial}{\partial b}V^{*}V^{T}\right) = \left(\frac{\partial}{\partial a}\left[j\omega\frac{\partial}{\partial b}(D_m-D_k)\ \exp\{j\omega(d_m-D_k)\}\right]\right),$$

$$= (j\omega)^2\left(\frac{\partial}{\partial b}(D_m-D_k)\frac{\partial}{\partial a}(D_m-D_k)\ \exp\{j\omega(D_m-D_k)\}\right)$$

$$+ (j\omega)\left(\frac{\partial}{\partial a}\ \frac{\partial}{\partial b}(D_m-D_k)\ \exp\{j\omega(D_m-D_k)\}\right)$$

$$= (j\omega)^2\ A + (j\omega)B\ , \qquad (2.29)$$

where the matrix A contains all and only those terms which involve differentiating the exponential twice.

Define $\tilde{A}$ and $\tilde{B}$ by

$$\tilde{A} = N^{-1}AN^{-1} \qquad (2.30)$$

and

$$\tilde{B} = N^{-1}BN^{-1} \qquad (2.31)$$

In view of the matrix identity $X^{T}QX^{*} = tr\ X^{*}X^{T}Q$, the expectation in Equation 2.27 is

$$< X^T N^{-1} \frac{\partial}{\partial a}(\frac{\partial}{\partial b} V^* V^T) N^{-1} X^* >$$

$$= < \text{tr } X^* X^T((j\omega)^2 \tilde{A} + (j\omega)\tilde{B}) >$$

$$= \text{tr}\{R[(j\omega)^2 \tilde{A} + (j\omega)\tilde{B})]\} \qquad (2.32)$$

$$= \text{tr}\{(N+SV^*V^T)((j\omega)^2 \tilde{A} + (j\omega)\tilde{B})\}.$$

The matrices A and B have only zeros for main diagonal elements. Thus $\tilde{A}$ and $\tilde{B}$ have only zeros for main diagonal elements, since N^{-1} is diagonal. Again, the main diagonal elements of $N\tilde{A}$ and $N\tilde{B}$ are all zeros. Thus $\text{tr}(N\tilde{A}) = \text{tr}(N\tilde{B}) = 0$, and

$$< X^T N^{-1} \frac{\partial}{\partial a}(\frac{\partial}{\partial b} V^* V^T) N^{-1} X^* >$$

$$= \text{tr}\{SV^*V^T((j\omega)^2 \tilde{A} + (j\omega)\tilde{B})\}. \qquad (2.33)$$

The k-th main diagonal element of the matrix

$$V^*V^T\tilde{B} = V^*V^T[(\frac{\partial}{\partial a}(\frac{\partial}{\partial b}(D_k - D_m)))\frac{1}{N_k N_m} \exp\{j\omega(D_k - D_m)\} \qquad (2.34)$$

is

$$(V^*V^T\tilde{B})_{kk} = \sum_m \exp\{j\omega(D_m - D_k)\}[\frac{\partial}{\partial a}(\frac{\partial}{\partial b}(D_k - D_m))]\frac{1}{N_k N_m} \exp\{j\omega(D_k - D_m)\}$$

$$= \sum_m \frac{1}{N_k N_m}[\frac{\partial}{\partial a}(\frac{\partial}{\partial b}(D_k - D_m))] \quad . \qquad (2.35)$$

It follows that

$$\text{tr}(V^*V^T\tilde{B}) = \sum\sum_{km} \frac{1}{N_k N_m}[\frac{\partial}{\partial a}(\frac{\partial}{\partial b}(D_k - D_m))]$$

$$= 0 \ , \qquad (2.36)$$

since

$$\frac{\partial}{\partial a}\left(\frac{\partial}{\partial b}(D_k - D_m)\right) = -\frac{\partial}{\partial a}\left(\frac{\partial}{\partial b}(D_i - D_m)\right) . \qquad (2.37)$$

Thus the only part of the expectation in the right side of Equation 2.27 that is not zero is the term involving $\mathrm{tr}\ V^*V^T\tilde{A}$.

The i-th main diagonal term of the matrix $V^*V^T\tilde{A}$ is

$$(V^*V^T\tilde{A})_{ii} = \sum_m \exp\{j\omega(D_m - D_i)\}\frac{\partial(D_i - D_m)}{\partial a}\,\frac{\partial(D_i - D_m)}{\partial b}\,\frac{\exp\{j\omega(D_i - D_m)\}}{N_i N_m}$$

$$= \sum_m \frac{\partial(D_i - D_m)}{\partial a}\,\frac{\partial(D_i - D_m)}{\partial b}\,\frac{1}{N_i N_m} . \qquad (2.38)$$

Using Equations 2.33, 2.36, and 2.38 in Equation 2.27, the typical element of the FIM becomes

$$-\left\langle \frac{\partial}{\partial a}\left(\frac{\partial}{\partial b}\ln L(Y)\right)\right\rangle = \sum_{B+} \omega^2 G \sum\sum_{km} \frac{S}{N_k N_m}\,\frac{\partial(D_k - D_m)}{\partial a}\,\frac{\partial(D_k - D_m)}{\partial b} . \qquad (2.39)$$

Since T is large and the spectra do not vary appreciably over intervals of width ω_o, Equation 2.39 can be written as

$$-\left\langle \frac{\partial}{\partial a}\left(\frac{\partial}{\partial b}\ln L(Y)\right)\right\rangle = \frac{T}{4\pi}\int_B d\omega\ \omega^2\,\frac{S^2}{\left(1 + \sum_i \frac{S}{N_i}\right)}$$

$$\cdot \sum\sum_{km} \frac{1}{N_k N_m}\,\frac{\partial(D_k - D_m)}{\partial a}\,\frac{\partial(D_k - D_m)}{\partial b} . \qquad (2.40)$$

For the likelihood function L(Y), the elements of the FIM as developed in Equation 2.39 and 2.40 do not depend on regarding a and b as elements of the delay vector D. The Equations 2.39 and 2.40 are correct so long as a and b are parameters, for example, range or bearing, which affect the likelihood function only in that a and b are arguments of the delays.

Specializing to the case where the gradient is taken with respect to the components of D, the MxM matrix of second partial derivatives is

$$- < \text{grad}(\text{grad } \ln L(D))^T > =$$

$$\sum_{B+} 2\omega^2 \frac{S^2}{1 + \sum_i \frac{S}{N_i}} \left[\left(\sum_i \frac{1}{N_i} \right) N^{-1} - N^{-1} \underline{1}\underline{1}^T N^{-1} \right] \qquad (2.41)$$

This is not the FIM, for it does not have rank M. This is due to the fact that the delays appear in the expression for L(D) only in difference pairs. One of the delays D_i is arbitrary. Arbitrarily set $D_1 = 0$. Redefine D by

$$D \equiv (D_2, D_3, \ldots, D_M) . \qquad (2.42)$$

The FIM pertinent to the estimation of the M-1 vector D defined in Equation 2.42 is obtained from Equation 2.41:

$$(\text{FIM}) = - < \text{grad}(\text{grad } \ln L(D))^T >$$

$$= \sum_{B+} 2\omega^2 \frac{S^2}{1 + \sum_i \frac{S}{N_i}} \left[\left(\sum_i \frac{1}{N_i} \right) N_P^{-1} - N_P^{-1} \underline{1}\, \underline{1}^T N_P^{-1} \right] \qquad (2.43)$$

In Equation 2.43 $N_P^{-1} = \text{diag}\,(N_2^{-1}, \ldots, N_M^{-1})$ and $\underline{1}$ is now an M-1 vector. Or equivalently, the FIM is the M-1 by M-1 lower right partition of the matrix on the right side of Equation 2.40.

An unbiased estimate $\hat{Y}$ of the vector Y is efficient if the covariance matrix of $Y-\hat{Y}$ is the CRMB. A necessary and sufficient condition for the existence of an efficient estimate is that

$$\text{grad ln } L(Y) = M(Y)(\hat{Y}-Y), \tag{2.44}$$

where M(Y) is a matrix which depends only on Y and not on the observations. (See Appendix A.) For the problem treated in this paper

$$\text{grad ln } (L(D) = \sum_{B+} \text{grad ln } X^T R^{-1} X^*. \tag{2.45}$$

Every term on the right side of Equation 2.45 has the observations, that is, the Fourier coefficients, as factors. Thus in general, efficient estimates do not exist. However, it is shown in the next two chapters that under certain conditions correlators and other square-law processing schemes are asymptotically efficient, in the limit of large T.

D. THE M=2 AND M=3 CASES

If the number of sensors is M = 2, the matrix bound reduces to a scalar bound. The scalar delay to be estimated is D, the delay from the first to the second sensor. The variance of any unbiased estimate, $\hat{D}$, of D satisfies

$$\mathrm{var}(\hat{D}) \geq \frac{2\pi}{T}\left[\int_B d\omega\ \omega^2\ \frac{S^2/N_1N_2}{1+\frac{S}{N_1}+\frac{S}{N_2}}\right]^{-1}. \qquad (2.46)$$

If the number of sensors is M = 3, and the vector to be estimated is defined as $Y^T = (Y_1,Y_2)^T = (D_2,D_3-D_2)^T$,

$$(\mathrm{FIM}) = \frac{T}{2\pi}\begin{bmatrix} K_{12}+K_{13} & K_{13} \\ K_{13} & K_{23}+K_{13}\end{bmatrix}, \qquad (2.47)$$

where

$$K_{ij} = \int_B \omega^2\ \frac{S^2/N_iN_j}{1+\sum_{k=1}^{3}\frac{S}{N_k}}\ d\omega \qquad (2.48)$$

It is shown in Chapter IV that for M = 3 a processing scheme using three properly filtered correlators achieves the matrix bound.

CHAPTER III

THE MAXIMUM LIKELIHOOD ESTIMATE

A. THE ASYMPTOTIC NATURE OF THE MLE

When the MLE is based on a large number of independent samples it is consistent, asymptotically normal, and asmptotically efficient [7]. Since this thesis treats the case in which the observation time T is large and the correlation times are small, there should be, in some sense, a large number of independent samples. It should, therefore, be possible to examine the errors in the MLE and show that the covariance matrix of the errors is in fact the CRMB. In this chapter this is demonstrated. Since the results developed in this chapter are independent of the true delays, the true delays are assumed to be all equal to zero. Without loss of generality, the signal delay at the first sensor is taken to be zero, and the delay set to be measured is the set of signal delays from the first to the remaining M-1 sensors.

The MLE for the delay vector, D, satisfies grad ln L(D) = 0, where the gradient is taken with respect to the M-1 unknown delays. The vector D so determined is the vector of measurement errors. The vector of delay measurement errors is

$$D^T = (D_2, D_3, \ldots, D_M) , \qquad (3.1)$$

and the error in the steering vector is

$$V^T = (1, \exp\{j\omega D_2\}, \ldots, \exp\{j\omega D_M\}). \qquad (3.2)$$

The error vectors D and V satisfy

$$\begin{aligned} 0 &= \text{grad } \ln L(D) \\ &= \text{grad} \sum_{B+} X^T R^{-1} X^* \\ &= \text{grad} \sum_{B+} G X^T N^{-1} V^* V^T N^{-1} X^* \\ &= \text{grad} \sum_{B+} G \sum_{in} \frac{X_i X_n^*}{N_i N_n} \exp\{j\omega(D_n - D_i)\} \qquad (3.3) \\ &= \text{grad} \sum_{B+} G \sum_{in} \frac{X_i X_n^*}{N_i N_n}(1 + j\omega(D_n - D_i) + \tfrac{1}{2}(j\omega)^2 (D_n - D_i)^2 + \text{HOT}) \\ &= \text{grad}(A + BD + \tfrac{1}{2} D^T C D + \text{HOT}). \end{aligned}$$

In Equation 3.3, $A + B^T D + \frac{1}{2} D^T C D + \text{HOT}$ is the series expansion, in terms of the vector D, for

$$\sum_{B+} G \sum_{in} (X_i X_n^* / N_i N_n) \exp\{j\omega(D_n - D_i)\}.$$

Since the error vector, D, is assumed to be small, Equation 3.3 leads to $0 = B + CD + \text{HOT}$, or to first order, $D = -C^{-1}B$. The remainder of this chapter justifies the following (asymptotic) Equations:

1) $\langle B^T \rangle = \underline{0}$ (3.4)

2) $\langle B^T B^* \rangle = (FIM)$ (3.5)

3) $C = \langle C \rangle = -(FIM)$ (3.6)

4) $\langle D \rangle = \underline{0}$ (3.7)

5) $\langle DD^T \rangle = (CRMB)$ (3.8)

B. THE VECTOR B

From Equation 3.3,

$$BD = \sum_{B+} G \sum_{i}\sum_{n} \frac{X_i X_n^*}{N_i N_n}(j\omega)(D_n - D_i)$$

$$= \sum_{B+} j\omega G \left[\left[\sum_i \frac{X_i}{N_i} \right] \left[\sum_n \frac{X_n^*}{N_n} D_n \right] - \left[\sum_n \frac{X_n^*}{N_n} \right] \left[\sum_i \frac{X_i}{N_i} D_i \right] \right]$$

$$= \left[\sum_{B+} j\omega G \underline{1}^T N^{-1} [X X_p^{*T} - X^* X_p^T] N_p^{-1} \right] D, \qquad (3.9)$$

where X_p is the X vector with the first component partitioned away. Thus

$$B = \sum_{B+} j\omega G \underline{1}^T N^{-1} [X X_p^{*T} - X^* X_p^T] N_p^{-1} . \qquad (3.10)$$

Since, at each frequency,

$$R \equiv < X^* X^T >$$

$$= N + S \underline{1}\underline{1}^T$$

$$= < X X^{*T} > , \qquad (3.11)$$

it follows that $< B > = 0$.

In examining the terms of $< B^T B^* >$ let

$$u_k \equiv \left(\sum_i \frac{X_i}{N_i} \right) \frac{X_k^*}{N_k} . \qquad (3.12)$$

Then

$$< u_k > = \sum_i \frac{R_{ik}}{N_i N_k}$$

$$= \frac{1}{N_k}\left(1 + \sum_i \frac{S}{N_i}\right)$$

$$= < u_k^* > . \qquad (3.13)$$

If ω_p and ω_q are distinct frequencies,

$$< (u_k(\omega_p)-u_k^*(\omega_p))(u_n(\omega_q)-u_n^*(\omega_q))^* >$$

$$= < (u_k(\omega_p)-u_k^*(\omega_p)) > < (u_n(\omega_q)-u_n^*(\omega_q)^* >$$

$$= 0 , \qquad (3.14)$$

since the Fourier coefficients at different frequencies are independent. The element in the (p-1)-th row and (q-1)-th column of $< B^T B^* >$ is, from Equations 3.9, 3.12, and 3.14,

$$< B^T B^* >_{p-1,q-1} = \sum_{B+} \omega^2 G^2 < (u_p-u_p^*)(u_q-u_q^*)^* > \qquad (3.15)$$
$$= \sum_{B+} \omega^2 G^2 < u_p u_q^* + u_p^* u_q - u_p^* u_q^* - u_p u_q > .$$

$$< u_p u_q^* > = \sum_{ik} \sum \frac{< X_i X_k^* X_p^* X_q >}{N_i N_k N_p N_q}$$

$$= \sum_{ik} \sum \frac{R_{ki} R_{pq} + R_{pi} R_{kq}}{N_i N_k N_p N_q} \qquad , \qquad (3.16)$$

since $< X_i X_q > = < X_k^* X_p^* > = 0$.

Since $R = N + S \underline{1}\,\underline{1}^T$ is symmetric,

$$< u_p^* u_q > = < u_p u_q^* > . \qquad (3.17)$$

Further,

$$< u_p u_q > = \sum_{ik} \sum \frac{< X_i X_k X_p^* X_q^* >}{N_i N_k N_p N_q}$$

$$= \sum_{ik} \sum \frac{R_{pi} R_{qk} + R_{pk} R_{qi}}{N_i N_k N_p N_q}$$

$$= < u_p^* u_q^* > . \qquad (3.18)$$

Combining the last four Equations, if follows that

$$< B^T B^* >_{p-1,q-1} = \sum_{B+} 2\omega^2 G^2 \sum\sum_{ik} \frac{R_{ik}R_{pq} - R_{pi}R_{qk}}{N_i N_k N_p N_q} . \qquad (3.19)$$

The double sum in Equation 3.19 is easily evaluated since

$$\sum\sum_{ik} \frac{R_{ik}R_{pq}}{N_i N_k N_p N_q} = \frac{R_{pq}}{N_p N_q} \sum \frac{1}{N_i} \left(\sum_k \frac{R_{ik}}{N_k} \right)$$

$$= \frac{R_{pq}}{N_p N_q} \sum \frac{1}{N_i} \left(1 + \sum_k \frac{S}{N_k}\right)$$

$$= \frac{R_{pq}}{N_p N_q} \left(\sum_i \frac{1}{N_i} \right) \left(1 + \sum_k \frac{S}{N_k}\right) , \qquad (3.20)$$

and

$$\sum\sum_{ik} \frac{R_{pi}R_{qk}}{N_i N_k N_p N_q} = \frac{1}{N_p N_q} \left(\sum_i \frac{R_{pi}}{N_i} \right) \left(\sum_k \frac{R_{qk}}{N_k} \right)$$

$$= \frac{1}{N_p N_q} \left(1 + \sum_i \frac{S}{N_i}\right)^2 . \qquad (3.21)$$

Thus Equation 3.19 becomes

$$< B^T B^* >_{p-1,q-1} = \sum_{B+} 2^2 G^2 \frac{R_{pq} \left(\sum_i \frac{1}{N_i} \right) - \left(1 + \sum_i \frac{S}{N_i}\right)}{N_p N_q} \left(1 + \sum_i \frac{S}{N_i}\right)$$

$$= \sum_{B+} 2\omega^2 \frac{S^2}{\left(1 + \sum_i \frac{S}{N_i}\right)} \left(\frac{\delta_{pq} N_p - 1}{N_p N_q} \right) . \qquad (3.22)$$

Examining the elements of $< B^T B^* >$ as given by Equation 3.22 shows that $< B^T B^* >$ is the M-1 by M-1 lower right partition of the matrix

$$F = \sum_{B+} 2\omega^2 \frac{S^2}{(1 + \sum_i \frac{S}{N_i})}[(\sum_i \frac{1}{N_i})N^{-1} - N^{-1} \underline{1}\underline{1}^T N^{-1}] \quad . \qquad (3.23)$$

But this lower right partition of F is the FIM for this problem, as was shown in Chapter II.

This completes the proof of Equations 3.4 and 3.5.

C. THE C MATRIX

From Equation 3.3

$$\frac{1}{2}D^T C D = \sum_{B+} \frac{1}{2}(j\omega)^2 G \sum\sum_{ik} \frac{X_i X_k^*}{N_i N_k}(D_k - D_i)^2 \ . \qquad (3.24)$$

The terms $X_i X_k^*$ are the elements of the sample covariance matrix based on T seconds of data. These sample covariance elements do not converge, even if T is arbitrarily long [8]. However, if the spectra being estimated are sufficiently smooth, the sample covariances can be averaged with samples from nearby frequencies to provide statistical convergence. By hypothesis, the spectra in the problem being treated are sufficiently smooth. The summation, $\sum_{B+}$, provides a smoothing over adjacent frequencies of the weighted spectral convariance estimates. Thus it is assumed that $C = \langle C \rangle$, and any statistical variations in CD are of second order relative to the variations in B. From Equation 3.3,

$$\frac{1}{2}D^T C D = \sum_{B+} \frac{1}{2}(j\omega)^2 G \sum\sum_{ik} \frac{X_i X_k^*}{N_i N_k}(D_k^2 + D_i^2 - 2 D_i D_k). \qquad (3.25)$$

The range of the indices i and k in the summations is from 1 to M. $D_1 = 0$. Using the argument preceding Equation 3.25,

$$\begin{aligned} \frac{1}{2}D^T C D &= \frac{1}{2}D^T \langle C \rangle D \\ &= \sum_{B+} (j\omega)^2 G \sum\sum_{ik} \frac{R_{ik}}{N_i N_k}(D_i^2 - D_i D_k) \ . \end{aligned} \qquad (3.26)$$

But

$$\sum_{ik}\sum \frac{R_{ik}}{N_i N_k} D_i^2 = \sum_i \frac{D_i^2}{N_i} \sum_k \frac{R_{ik}}{N_k}$$

$$= \sum_i \frac{D_i^2}{N_i} \left[1 + \sum_k \frac{S}{N_k}\right]$$

$$= \left[1 + \sum_k \frac{S}{N_k}\right] D^T N_P^{-1} D \ , \qquad (3.27)$$

where N_P^{-1} is the N^{-1} matrix with the first row and column partitioned away. In similar manner

$$\sum_{ik}\sum \frac{R_{ik}}{N_i N_k} D_i D_k = D^T N_P^{-1} R_P N_P^{-1} D \ , \qquad (3.28)$$

where R_P is the appropriately partitioned R matrix. Thus

$$\frac{1}{2} D^T < C > D = \sum_{B+} (j\omega)^2 G\{(1 + \sum_i \frac{S}{N_i}) D^T N_P^{-1} D - D^T N_P^{-1} R_P N_P^{-1} D\}$$

$$= D^T\{ \sum_{B+} (j\omega)^2 G[(\sum_i \frac{S}{N_i}) N_P^{-1} - S N_P^{-1} \underline{1}\underline{1}^T N_P^{-1}]\} D. \qquad (3.29)$$

It follows that

$$< C > = \sum_{B+} 2(j\omega)^2 \frac{S^2}{1 + \sum_i \frac{S}{N_i}} \left[(\sum_i \frac{1}{N_i}) N_P^{-1} - N_P^{-1} \underline{1}\underline{1}^T N_P^{-1}\right] . \qquad (3.30)$$

In fact,

$$- < C > = < B^T B^* > = (FIM) \ , \qquad (3.31)$$

which demonstrates Equation 3.6.

D. THE D VECTOR

Since the measurement errors are small, the D vector is essentially determined by $\underline{0} = B^T + C^T D$. Since $C = \langle C \rangle$ with negligible error, and since $\langle C \rangle$ is symmetric, it follows that, asymptotically,

$$D = - \langle C \rangle^{-1} B^T , \tag{3.32}$$

$$\langle D \rangle = - \langle C \rangle^{-1} \langle B^T \rangle = \underline{0} , \tag{3.33}$$

and the covariance matrix for D is

$$\begin{aligned} \langle D^* D^T \rangle &= \langle C \rangle^{-1} \langle B^T B^* \rangle^* \langle C \rangle^{-1} \\ &= \langle C \rangle^{-1} (\text{FIM}) \langle C \rangle^{-1} \\ &= (\text{FIM})^{-1} \\ &= (\text{CRMB}) . \end{aligned} \tag{3.34}$$

Thus the MLE for the delay vector is efficient. It is also unbiased. This demonstrates Equations 3.7 and 3.8.

CHAPTER IV

CORRELATOR DELAY MEASUREMENT SYSTEMS

In this chapter a delay measurement system using correlator techniques is studied. The covariance matrix for the delay estimates is derived. The use of filters to optimize the system is studied. For the M = 2 and M = 3 cases, the correlator based measurement system with certain filters is shown to achieve the CRMB. A conjecture is made regarding the optimum filters when M > 3.

A. CORRELATOR MEASUREMENT COVARIANCES

The output of an ideal correlator with T seconds of integration, and with the i-th and j-th sensors for inputs, is

$$z_{ij}(T_j) = \frac{1}{T} \int_T x_i(t) x_j(t - T_j)\, dt \ . \qquad (4.1)$$

By the assumptions in Chapter II, the additive sensor noises are independent. It follows then, that

$$< z_{ij}(T_j) > = R_s(D_i - D_j - T_j) \ , \qquad (4.2)$$

and that the peak of the function $z_{ij}(T_j)$ determines an unbiased estimate for $D_i - D_j$.

Let $x_i(t)$, $x_j(t)$ and $x_k(t)$, $x_n(t)$ be two pairs of sensor outputs. These pairs are correlated to determine $z_{ij}(T_j)$ and $z_{kn}(T_n)$. The time records $x_i(t)$ and $x_j(t)$ are assumed to be the outputs from two distinct sensors. The same applies to $x_k(t)$ and $x_n(t)$. However, it is not assumed that all four inputs are distinct. With the integers i, j, k, and n fixed, it is convenient, for this section, to relabel $x_i(t)$ as $x_1(t)$, $x_j(t)$ as $x_2(t)$, $x_k(t)$ as $x_3(t)$, and $x_n(t)$ as $x_4(t)$. Thus, the $x_1(t)$ of this section may or may not be the $x_1(t)$ of Chapter II. And the $x_1(t)$ of this section may be, in fact, the same as $x_3(t)$ or $x_4(t)$.

The correlators have as outputs $z_{12}(T_2)$ and $z_{34}(T_4)$. Define $y_{12}(T_2)$ and $y_{34}(T_4)$ by:

$$y_{12}(T_2) = z_{12}(T_2) - \langle z_{12}(T_2) \rangle$$

$$y_{34}(T_4) = z_{34}(T_4) - \langle z_{34}(T_4) \rangle \qquad (4.3)$$

Let $\hat{T}_2$ and $\hat{T}_4$ be those values of T_2 and T_4, respectively, at which the correlogram peaks are located. Let e_2 and e_4 be defined by:

$$e_2 + T_2 = D_1 - D_2$$

$$e_4 + T_4 = D_3 - D_4 \qquad (4.4)$$

and let $\hat{e}_2$ and $\hat{e}_4$ be, respectively, the values of e_2 and e_4 when $T_2 = \hat{T}_2$ and $T_4 = \hat{T}_4$. Then $\hat{e}_2$ and $\hat{e}_4$ are the measurement errors. The correlator outputs are with probability 1 everywhere differentiable functions of T_2 and T_4. $\hat{T}_2$ and $\hat{T}_4$ are determined by:

$$\frac{d}{dT_2} z_{12}(T_2) = 0 \qquad \text{at } T_2 = \hat{T}_2$$

$$\frac{d}{dT_4} z_{34}(T_4) = 0 \qquad \text{at } T_4 = \hat{T}_4 \qquad (4.5)$$

Letting the derivatives be denoted by primes, and using the definitions, Equations (4.5) become:

$$0 = R_s(\hat{e}_2) + y'_{12}(\hat{T}_2)$$

$$0 = R_s(\hat{e}_4) + y'_{34}(\hat{T}_4) \qquad (4.6)$$

The errors are assumed small. Assuming that $R_s(e)$ is differentiable at e = 0, it has the series expansion

$$R_s(e) = R_s(0) + \frac{1}{2} R_s''(0)e^2 + \text{HOT} . \qquad (4.7)$$

Using Equation 4.7 and neglecting the HOT, Equations 4.6 yield:

$$\hat{e}_2 = - (R_s''(0))^{-1} y_{12}(\hat{T}_2)$$

$$\hat{e}_4 = - (R_s''(0))^{-1} y_{34}(\hat{T}_4) , \qquad (4.8)$$

whence

$$< \hat{e}_2\hat{e}_4 > = (R_s''(0))^{-2} < y_{12}'(\hat{T}_2)y_{34}'(\hat{T}_4) > \qquad (4.9)$$

Next, the expectation on the right side of Equation 4.9 is evaluated. The evaluation is accomplished by using Fourier coefficients and Fourier series. The coefficients are as defined and used in Chapter II, and the summations that follow are understood to range over the (positive and negative) integers.

$$x_1(t) = \frac{1}{T} \sum_k X_1(k) \exp\{jk\omega_o t\} . \qquad (4.10)$$

$$x_2(t - T_2) = \frac{1}{T} \sum_k X_2^*(k) \exp\{-jk\omega_o (t-T_2)\} . \qquad (4.11)$$

$$y_{12}(T_2) = \frac{1}{T}\int_T x_1(t)x_2(t-T_2)\,dt - \langle x_1(t)x_2(t-T_2)\rangle$$

$$= \sum_{ik}\sum \left(\frac{1}{T^3}\int_T dt\,\exp\{j\omega_o t(i-k)\}\right)X_1(i)X_2^*(k)\,\exp\{jk\omega_o T_2\}$$

$$- \langle x_1(t)x_2(t-T_2)\rangle$$

$$= \frac{1}{T^2}\sum_k X_1(k)X_2^*(k)\,\exp\{jk\omega_o T_2\}$$

$$- \langle x_1(t)x_2(t-T_2)\rangle\,, \qquad (4.12)$$

since

$$\int_T dt\,\exp\{j\omega_o t(i-k)\} = T\,\delta_{ik}. \qquad (4.13)$$

$$y'_{12}(T_2) = \frac{1}{T^2}\sum_k jk\omega_o X_1(k)X_2^*(k)\,\exp\{jk\omega_o T_2\}$$

$$- \frac{d}{dT_2}\langle x_1(t)x_2(t-T_2)\rangle\,. \qquad (4.14)$$

From Equation 4.14, $\langle y'_{12}(T_2)\rangle = 0$, since both terms on the right side of Equation 4.14 are, after the expectation is taken, equal to $\pm R'_s(D_1 - D_2 - T_2)$.

An expression for $y'_{34}(T_4)$ can be written by replacing 1 and 2 where they occur in Equation 4.14 by 3 and 4, respectively. Thus

$$< y'_{12}(T_2)y'_{34}(T_4) > =$$

$$\frac{1}{T^4} \sum_{kn}\sum [(j\omega_o k)(j\omega_o n) < X_1(k)X_2^*(k)X_3(n)X_4^*(n) >$$

$$\cdot \exp\{j\omega_o(kT_2 + nT_4)\}]$$

$$- R'_s(D_1-D_2-T_2)R'_s(D_3-D_4-T_4).$$

$$= \frac{1}{T^4} \sum_{kn}\sum (j\omega_o k)(j\omega_o n)[< X_1(k)X_3(n) > < X_2^*(k)X_4^*(n) >$$

$$+ < X_1(k)X_4^*(n) > < X_3(n)X_2^*(k) >] . \qquad (4.15)$$

The last part of Equation 4.15 follows from the argument made immediately after Equation 4.14.

The Fourier coefficients at positive and negative frequencies are related in such a way that

$$< X_1(k)X_3(n) > = < X_1(k)X_3^*(-n) >$$

$$= \begin{cases} < X_1(k)X_3^*(k) > & \text{if } n = -k, \\ 0 & \text{if } n \neq -k. \end{cases} \qquad (4.16)$$

Thus,

$$< y'_{12}(T_2)y'_{34}(T_4) >$$

$$= \frac{1}{T^4} \sum_{n} (n\omega_o)^2[< X_1X_3^* > < X_2^*X_4 > \exp\{jn\omega_o(T_2-T_4)\}$$

$$- < X_1X_4^* > < X_2X_2^* > \exp\{jn\omega_o(T_2+T_4)\}] . \qquad (4.17)$$

Now replace 1, 2, 3, and 4 everywhere in Equation 4.17 by h, i, k, and l, respectively. Note that

$$< X_h X_i^* > = T(N_h\delta_{hi} + S \exp\{j\omega(D_i - D_h)\})$$

$$= T(S + N_h\delta_{hi}) \exp\{j\omega(D_i - D_h)\} . \qquad (4.18)$$

Hence Equation 4.17 becomes

$$< y'_{hi}(T_i) y'_{kl}(T_l) > = \frac{1}{T^2} \sum_n (n\omega_o)^2$$

$$\cdot[(S + N_h\delta_{hk})(S + N_i\delta_{il}) \exp\{jn\omega_o A\}$$

$$-(S + N_h\delta_{hl})(S + N_i\delta_{ik}) \exp\{jn\omega_o B\}], \qquad (4.19)$$

where

$$A = T_i - T_l - D_h + D_k + D_i - D_l , \qquad (4.20)$$

and

$$B = T_i + T_l - D_h + D_l - D_k + D_i . \qquad (4.21)$$

Next let $T_i = \hat{T}_i$ and $T_l = \hat{T}_l$. Then A and B become

$$\hat{A} = \hat{e}_i - \hat{e}_l \qquad (4.22)$$

and

$$\hat{B} = \hat{e}_i + \hat{e}_l , \qquad (4.23)$$

respectively. Neglecting the HOT in the errors, Equation 4.9 can be written as

$$< \hat{e}_i \hat{e}_l > = (R_s''(0))^{-2} \frac{1}{2\pi T} \int_B d\omega\ \omega^2 \qquad (4.24)$$

$$\cdot\{(S + N_h\delta_{hk})(S + N_i\delta_{il}) - (S + N_h\delta_{hl})(S + N_i\delta_{ik})\}.$$

The elements of the covariance matrix of the correlator delay estimates are determined by Equation 4.24.

B. FILTERS FOR CORRELATOR SYSTEMS (M=2 AND 3)

In this section the use of filters to optimize the corelator delay measurements is studied. For the M = 2 and M = 3 cases the optimal filters are displayed. For M > 3, the optimal filters are conjectured.

Consider first the M = 2 case. There is only one delay to measure. Let it be $D = D_2 - D_1$. The estimate for D is denoted $\hat{D}$. From Equation 4.24 and 2.46,

$$\text{var}(\hat{D}) = \frac{2\pi}{T} \frac{\int_B \omega^2 [N_1 N_2 + S(N_1 + N_2)]\, d\omega}{\left(\int_B \omega^2 S\, d\omega \right)^2}$$

$$\geq \frac{2\pi}{T} \left[\int_B \omega^2 \frac{S^2/N_1 N_2}{1 + \frac{S}{N_1} + \frac{S}{N_2}}\, d\omega \right]^{-1} . \qquad (4.25)$$

If filters $F_1(\omega)$ and $F_2(\omega)$ are to be used on the outputs of the sensors 1 and 2, respectively, the filters must have identical phase responses. If they do not, then the filtered signals will have different delays at different frequencies, and this could bias the delay estimate. Since the phase responses are identical, the filters may be assumed to be identical filters. Let $F = F_1 = F_2$. The filtered signal spectrum is $S(\omega)|F(\omega)|^2$. Similarly, the filtered noise spectra are $N_1(\omega)|F(\omega)|^2$ and $N_2(\omega)|F(\omega)|^2$. Thus Equation 4.25 is readily modified to account for the filtering. Equation 4.25 becomes

$$\mathrm{var}(\hat{D}) = \frac{2\pi}{T} \frac{\int_B \omega^2 |F|^4 [N_1 N_2 + S(N_1 + N_2)]\, d\omega}{\left(\int_B \omega^2 |F|^2 S\, d\omega \right)^2}$$

$$\geq \frac{2\pi}{T} \left[\int_B \omega^2 \frac{S^2/N_1 N_2}{1 + \frac{S}{N_1} + \frac{S}{N_2}}\, d\omega \right]^{-1}. \qquad (4.26)$$

If the filter $|F|^2$ is defined by

$$|F(\omega)|^2 = \frac{S/N_1 N_2}{1 + \frac{S}{N_1} + \frac{S}{N_2}} \quad , \qquad (4.27)$$

Equation 4.26 becomes an equality. The measurement variance is now the Cramér-Rao bound. Thus, although efficient estimates were shown in Chapter II not to exist for this case, none the less, a correlator system with filters determined by Equation 4.27 provides an asymptotically efficient delay estimate.

For the M > 2 case, there are M(M-1)/2 sensor pairs to be correlated. Let e_{ij} be the error in the delay estimate based on the correlation,

$$z_{ij}(T_j) = \frac{1}{T} \int_T x_i^F(t) x_j^F(t-T_j)\, dt \quad , \qquad (4.28)$$

where $x_i^F(t)$ and $x_j^F(t)$ are the filtered $x_i(t)$ and $x_j(t)$ processes, respectively. From Equation 4.24

$$\langle \hat{e}_{ij} \hat{e}_{kl} \rangle = \frac{2\pi}{T} \frac{\int_B \omega^2 |F_{ij}|^2 |F_{kl}|^2 G(ij;kl)\, d\omega}{\int_B \omega^2 |F_{ij}|^2 S\, d\omega \int_B \omega^2 |F_{kl}|^2 S\, d\omega} \quad , \qquad (4.29)$$

where

$$G(ij;kl) = [(S+N_i\delta_{ik})(S+N_j\delta_{jl}) - (S+N_i\delta_{il})(S+N_j\delta_{jk})] \ . \tag{4.30}$$

In Equation 4.29, $|F_{ij}(\omega)|$ is the filter used on the x_i and x_j waveforms, which are the inputs to the correlator that computes $z_{ij}(T_j)$.

Let the M(M-1)/2 errors be displayed as a column vector e, and let P_e be the covariance matrix for the errors:

$$P_e = < ee^T > \ . \tag{4.31}$$

The main diagonal terms of P_e are the individual variances of the M(M-1)/2 delay measurements from the M(M-1)/2 correlators. From the discussion pertaining to Equation 4.26, the filters

$$|F_{ij}|^2 = \frac{S/N_iN_j}{1 + \frac{S}{N_i} + \frac{S}{N_j}} \tag{4.32}$$

will minimize the main diagonal terms of P_e. It is tempting to conclude that these filters will then be the optimal filters. But this is not true in view of what follows.

For the problem being considered, a vector (of delays) is to be optimally estimated. The observations (from the correlators) are a linear combination of the delays to be estimated, corrupted by additive zero mean noise. Therefore, the vector to be estimated, D, is linearly related to the observations, T, by

$$T = AD + e \ . \tag{4.33}$$

The covariance matrix for e, P_e, is given in Equation 4.31. It is known that the optimal linear estimate for the vector D is the Gauss-Markov estimate [9],

$$\hat{D} = [A^T P_e^{-1} A]^{-1} A^T P_e^{-1} T \ . \tag{4.34}$$

The Gauss-Markov estimate is the minimum variance linear unbiased estimate for estimating a vector, given linear observations corrupted by additive zero-mean noise. The covariance matrix for the Gauss-Markov estimate is

$$< (\hat{D}-D)(\hat{D}-D)^T > = [A^T P_e^{-1} A]^{-1} \ . \tag{4.35}$$

Without loss of generality, it is convenient to fix the order of the elements, T_{ij}, of T by defining

$$T^T = (T_{12}, T_{13}, \ldots, T_{(M-1)M}) \ . \tag{4.36}$$

In addition define the following scalars, vectors, and matrices:

$$e^T = (e_{12}, e_{13}, \ldots, e_{(M-1)M}) \tag{4.37}$$

$$F = \text{diag}\ (|F_{12}|^2, |F_{13}|^2, \ldots, |F_{(M-1)M}|^2) \tag{4.38}$$

$$K = \int_B \omega^2 SF \, d\omega$$

$$= \text{diag}\ (K_{12}, K_{13}, \ldots, K_{(M-1)M}) \tag{4.39}$$

$$G = [G(ij;kl)] \ . \tag{4.40}$$

The matrix G has for its elements the terms G(ij;kl), positioned according to the scheme determined by the order of the subscripts in Equation 4.36. Then from Equations 4.29, 4.31, and 4.35 through 4.40,

$$P_e = \frac{2\pi}{T} K^{-1}\left(\int_B \omega^2 \, FGF \, d\omega\right)K^{-1} , \qquad (4.41)$$

and

$$< (\hat{D}-D)(\hat{D}-D)^T >^{-1} = \frac{T}{2\pi} A^T K\left(\int_B \omega^2 FGF \, d\omega\right)^{-1} KA . \qquad (4.42)$$

Consider the M = 3 case. The matrix G is

$$G = \begin{bmatrix} N_1N_2 + S(N_1+N_2) & SN_1 & -SN_2 \\ SN_1 & N_1N_3 + S(N_1+N_3) & SN_3 \\ -SN_2 & SN_3 & N_2N_3 + S(N_2+N_3) \end{bmatrix} . \qquad (4.43)$$

Let the filters be specified by

$$|F_{ij}|^2 = \frac{S/N_iN_j}{1 + \sum_{k=1}^{3} \frac{S}{N_k}} . \qquad (4.44)$$

Then

$$FGF = \frac{S^2}{1 + \sum_k \frac{S}{N_k}} \operatorname{diag}\left(\frac{1}{N_1N_2}, \frac{1}{N_1N_3}, \frac{1}{N_2N_3}\right)$$

$$- \frac{S^3/N_1N_2N_3}{(1 + \sum_k \frac{S}{N_k})^2} \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix} . \qquad (4.45)$$

Since

$$\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \end{bmatrix} , \tag{4.46}$$

define $w^T = (1 \quad -1 \quad 1)$, so that

$$P_e = \frac{2\pi}{T} K^{-1}[K + bww^T]K^{-1} , \tag{4.47}$$

where

$$b = - \int_B \omega^2 \frac{S^3/N_1N_2N_3}{(1 + \sum_k \frac{S}{N_k})^2} d\omega . \tag{4.48}$$

By the matrix inverse lemma [7]

$$P_e^{-1} = \frac{T}{2\pi} [K - \frac{ww^T}{\frac{1}{b} + w^TKw}] \tag{4.49}$$

$$= \frac{T}{2\pi} \frac{1}{d}[dK - ww^T] , \tag{4.50}$$

with

$$d = \frac{1}{b} + w^TKw . \tag{4.51}$$

It is now necessary to specify the vector D and the matrix A. Let D be defined as

$$D = (D_2, D_3 - D_2) , \tag{4.52}$$

so that

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} . \tag{4.53}$$

Then

$$< (\hat{D}-D)(\hat{D}-D)^T >^{-1} = A^T P_e^{-1} A$$

$$= \frac{T}{2\pi d} A^T \begin{bmatrix} dK_{12}-1 & 1 & -1 \\ 1 & dK_{13}-1 & 1 \\ -1 & 1 & dK_{23}-1 \end{bmatrix} A$$

$$= \frac{T}{2\pi d} \begin{bmatrix} d(K_{12}+K_{13}) & [illegible] \\ dK_{13} & d(K_{[illegible]}+K_{23}) \end{bmatrix} . \quad (4.54)$$

For the specific choice of filters,

$$K_{ij} = \int_B \omega^2 \frac{S^2/N_i N_j}{1 + \sum_k \frac{S}{N_k}} d\omega \quad . \quad (4.55)$$

The FIM matrix for M = 3 with the same choice of delay vector to be estimated was given in Equations 2.47 and 2.48. Those Equations are the same as Equations 4.54 and 4.55. That is, the Gauss-Markov estimate obtained from the three correlator processing scheme, with the inputs to the correlators filtered according to Equation 4.44, achieves the CRMB.

Thus, for M = 2 and M = 3, the correlator processing scheme can provide the best possible estimate, if the proper filters are used.

It is conjectured that the optimum filters to use with the correlator delay measurement scheme are defined for all M by

$$|F_{ij}|^2 = \frac{S^2/N_i N_j}{1 + \sum_k \frac{S}{N_k}} \quad . \tag{4.56}$$

CHAPTER V

SUBOPTIMAL FILTERING

In the preceding chapters optimum filtering schemes for correlator delay measurement systems were discussed. An interesting question is how sensitive are the measurement errors to changes in the filter design from the optimum. These design changes may be deliberate or inadvertant. An instance of the former is when the prior knowledge of the signal spectrum is limited. In such a case the designer may choose to design for an assumed worst case. Or perhaps he may choose to simply whiten the input noise Or perhaps the correlator system is also to be used for detection. The designer may choose to use the Eckart filters, since they are the optimal filters for correlator detectors [10].

In this chapter, the degradation in the correlator estimator performance is studied for M = 2. It is assumed that both noises have the same spectrum. The signal and noise spectra, S(ω) and N(ω), respectively, are both taken to be bandlimited with constant slopes of 0, -3, or -6 dB/octave.

$$S(\omega) = \begin{cases} 2A\ (\omega/\omega_1)^{-a}, & 1 \le \omega/\omega_1 \le 1 + W \\ 0, & \text{elsewhere}. \end{cases} \qquad (5.1)$$

$$N(\omega) = \begin{cases} 2B\ (\omega/\omega_1)^{-b}\ , & 1 \leq \omega/\omega_1 \leq 1 + W \\ 0\ , & \text{elsewhere}\ . \end{cases} \qquad (5.2)$$

In Equations 5.1 and 5.2 $S(\omega)$ and $N(\omega)$ are the one sided spectra. W is a bandwidth variable. For W = 1, 3, 7, and 15, the bandwidths of the spectra are 1, 2, 3, and 4 octaves, respectively. The spectrum slope is determined by a and b, each of which will be equal to 0, 1, or 2. The suboptimal filters will be defined by:

$$(1) \qquad |F_{NF}|^2 = 1 \qquad (5.3)$$

$$(2) \qquad |F_{ECK}|^2 = S/N^2 \qquad (5.4)$$

$$(3) \qquad |F_{WH}|^2 = 1/N \qquad (5.5)$$

$$(4) \qquad |F_{OPT}|^2 = \frac{S/N^2}{1 + 2\frac{S}{N}} \qquad (5.6)$$

F_{ECK} is the Eckart filter. F_{WH} is the filter that whitens the input noise. F_{OPT} is the optimal filter for delay estimation. Notice that for large SNRs (Signal to Noise Ratios) the optimal filter essentially whitens the noise. (The system performance is unaffected by filter gain constants.) For small SNRs the optimal filter is essentially an Eckart filter.

In what follows the system dependence on the input SNR is studied by plotting curves for five SNRs determined by

$$\left(\frac{A}{B}\right) = 2^{-k} \quad , \; k = 0, 1, 2, 3, 4. \tag{5.7}$$

The ordinates of the curves will be the measurement standard deviations for the optimally filtered case, or the degradation in dB when suboptimal filters are used. The variable W is used for the abscissa.

Whelchel [11] used the spectra of (5.1) and (5.2) in an analysis of the effects of suboptimal filtering on correlator performance. Whelchel was primarily interested in signal detection, and used the output SNR of the correlator as a measure of detection capability. The study of suboptimally filtered correlator estimators in this chapter in part parallels Whelchel's study of suboptimally filtered correlator detectors.

Denote the measurement variances by D^2_{NF}, D^2_{ECK}, D^2_{WH}, and D^2_{OPT}, to correspond to the filters F_{NF}, F_{ECK}, F_{WH}, and F_{OPT}.

Let $d = a - b$. When the spectra of (5.1) and (5.2) are used with the filters defined by (5.3), (5.4), (5.5), and (5.6), and the estimation variance is calculated using (4.26), the result in all cases except one depends on d, and not specifically on a and b. Only when the filters F_{WH} are used does the corresponding variance depend specifically on both a and b.

Graphs showing how the standard deviation of the optimally filtered delay estimate depends on the parameters

k and W are given in Figures 1, 2, and 3 for d = 0, 1, and 2, respectively. The curves are normalized by the factor $\sqrt{(\omega_1^3\ T)}$, so that $\sqrt{(\omega_1^3\ T)}\cdot\ D_{OPT}(W)$ is dimensionless. If for example, w_1 and T are chosen so that $\sqrt{(\omega_1^3\ T)} = 10^3 sec^{-1}$, then the abscissas of these three figures read directly in milliseconds. For d = 0 and d = 1 $D_{OPT}(W)$ can be made arbitrarily small by letting W be sufficiently large. This is not true for d = 2. The curves of Figure 1 also apply to D_{WH} and D_{ECK}, since for d = 0, S/N is simply a constant. They also apply to D_{NF} when the noise and signal spectra are both flat (a = b = C).

The asymptotic nature of the curves for small values of W may not truly represent the behavior of the system measurement error. This is because the derivation leading to (4.25) assumes a sufficiently large time-bandwidth product to yield measurements with small errors.

Figures 4 through 7 show the processor performance (in dB, relative to the optimal) when suboptimum filters D_{WH} and D_{ECK} are used. In these curves both the optimum and the suboptimum systems are presumed to process the same band of frequencies, determined by the argument W. For example, in Figure 4 the system degradation is given by

$$\text{dB loss} = 10 \log_{10} (D^2_{ECK}\ (W)/D^2_{OPT}(W)), \qquad (5.8)$$

for d = a - b = 1. As indicated by Figures 5 and 6, for the

assumed spectra, the Eckart filtered system sustains a processing loss relative to the optimum of at most about 0.5 dB. The processing loss decreases as the input SNR decreases, as expected.

Figures 8 and 9 depict the system performance if only flat band-pass filters, F_{NF}, are used. For these figures d = 2, with a = 2 and b = 0 in Figure 8, and a = 3 and b = 1 in Figure 9. In these figures the system degradation is measured relative to the infinite bandwidth ($W = \infty$) optimally filtered system. Thus the

$$\text{dB loss} = 10 \log_{10} \left(D^2_{NF}(W)/D^2_{OPT}(\infty)\right) , \qquad (5.9)$$

in Figures 8 and 9. Note that Figure 8 also gives the dB loss of $D^2_{WH}(W)$ relative to $D^2_{OPT}(\infty)$ for d = 2. This is because b = 0 in Figure 8, and the noise is flat. In both figures the processing loss at first decreases to a minimum, and then as too much high frequency noise is processed the variance then increases. Since the signal spectrum is falling off at 6 dB/octave faster than the noise spectrum, the processor will behave ever more poorly as the processor bandwidth is made larger and larger. It is interesting that a W of 3 or 4 corresponding to about a two octave processing band yields a loss of only 3 or 4 dB.

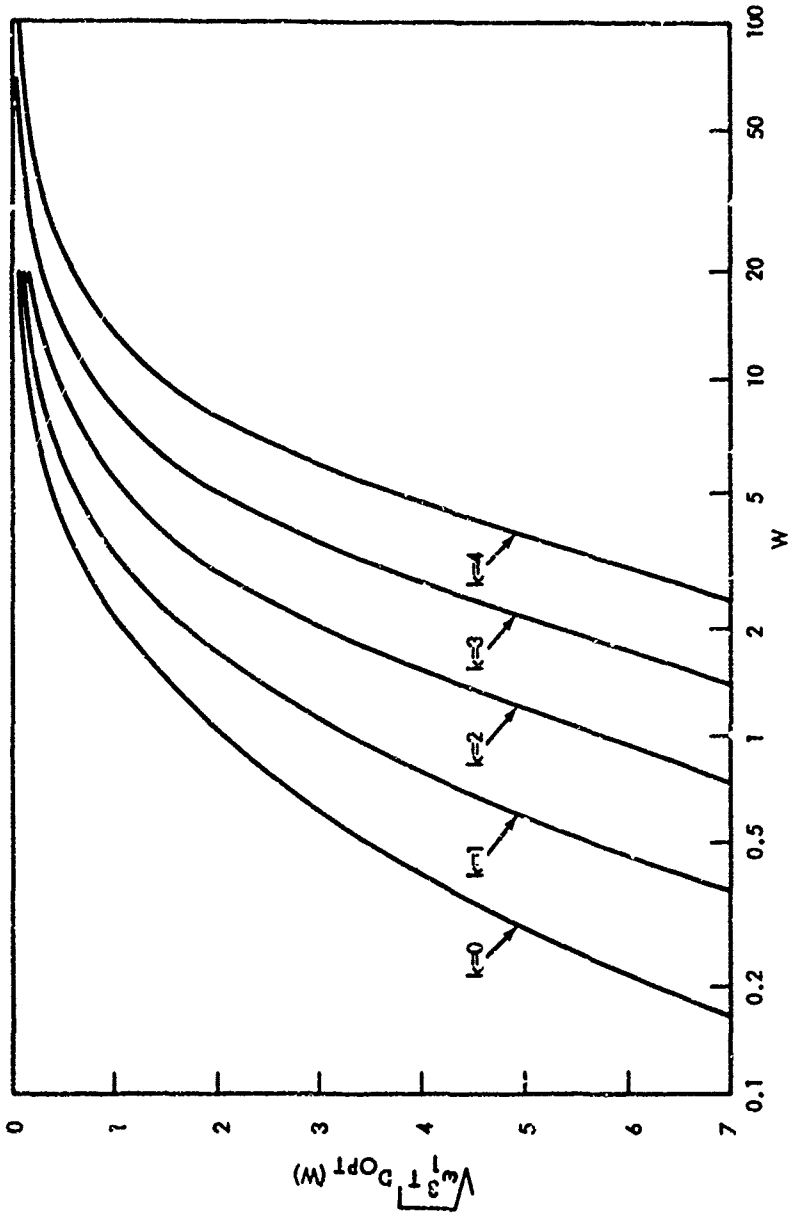

FIG. 1 NORMALIZED STANDARD DEVIATION, OPTIMUM, d = 0.

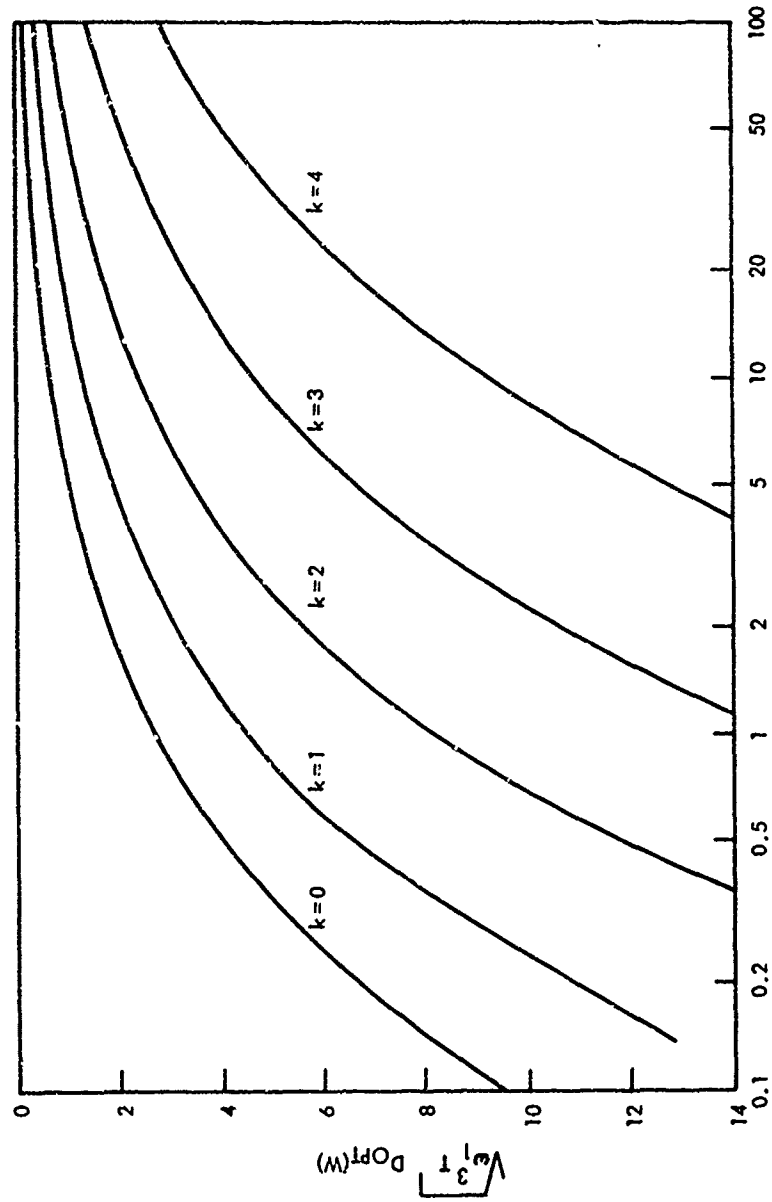

FIG. 2 NORMALIZED STANDARD DEVIATION, OPTIMUM, d=1.

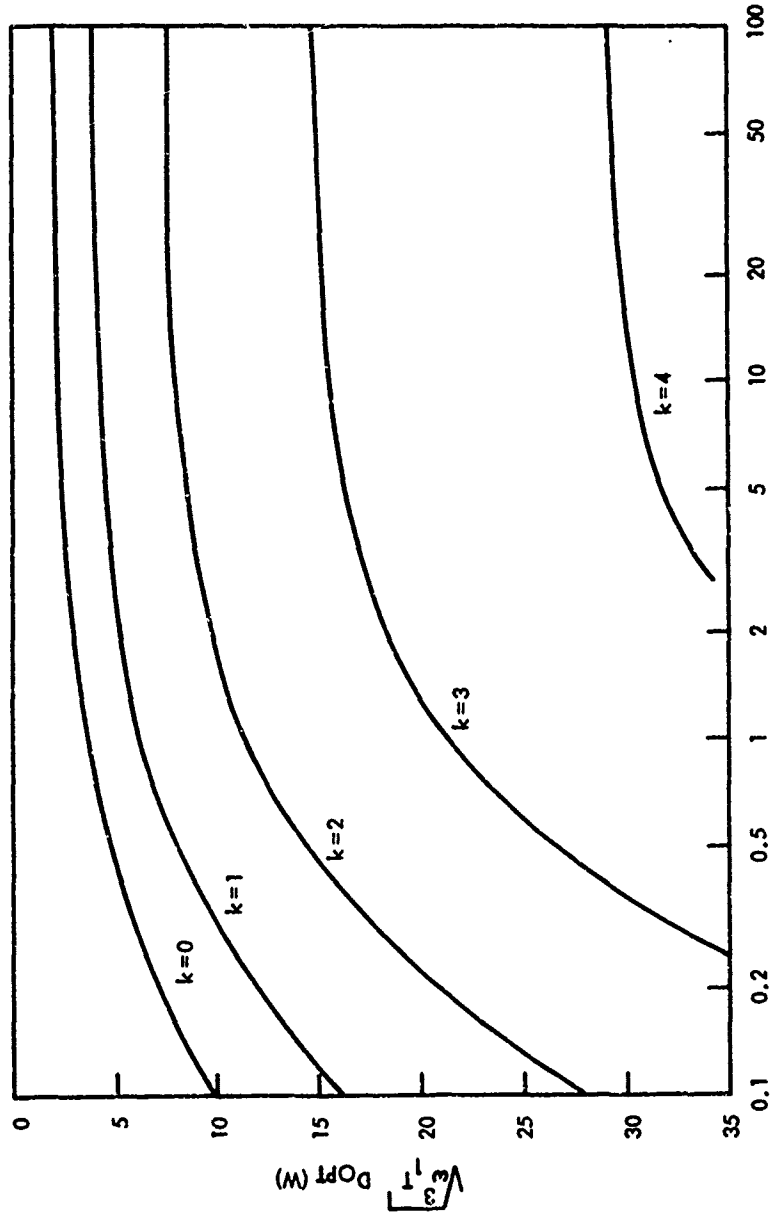

FIG. 3 NORMALIZED STANDARD DEVIATION, OPTIMUM, d = 2.

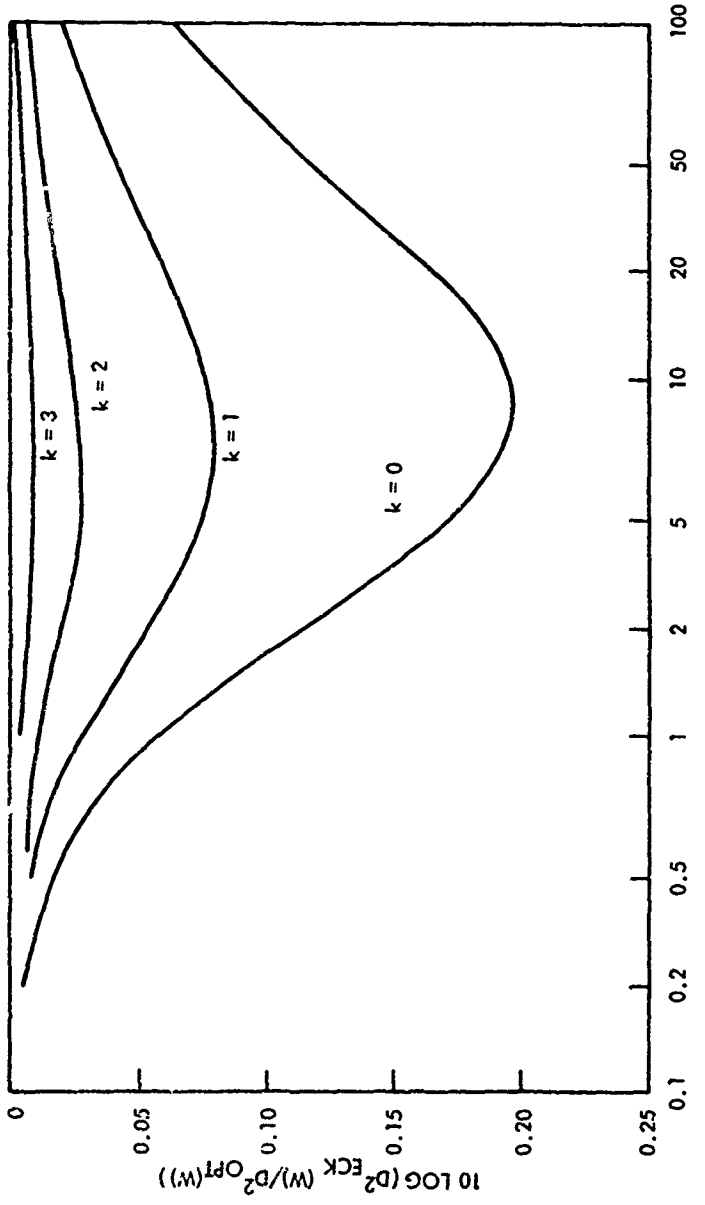

FIG. 4 SYSTEM DEGRADATION USING $|F_{ECK}|^2$, d=1.

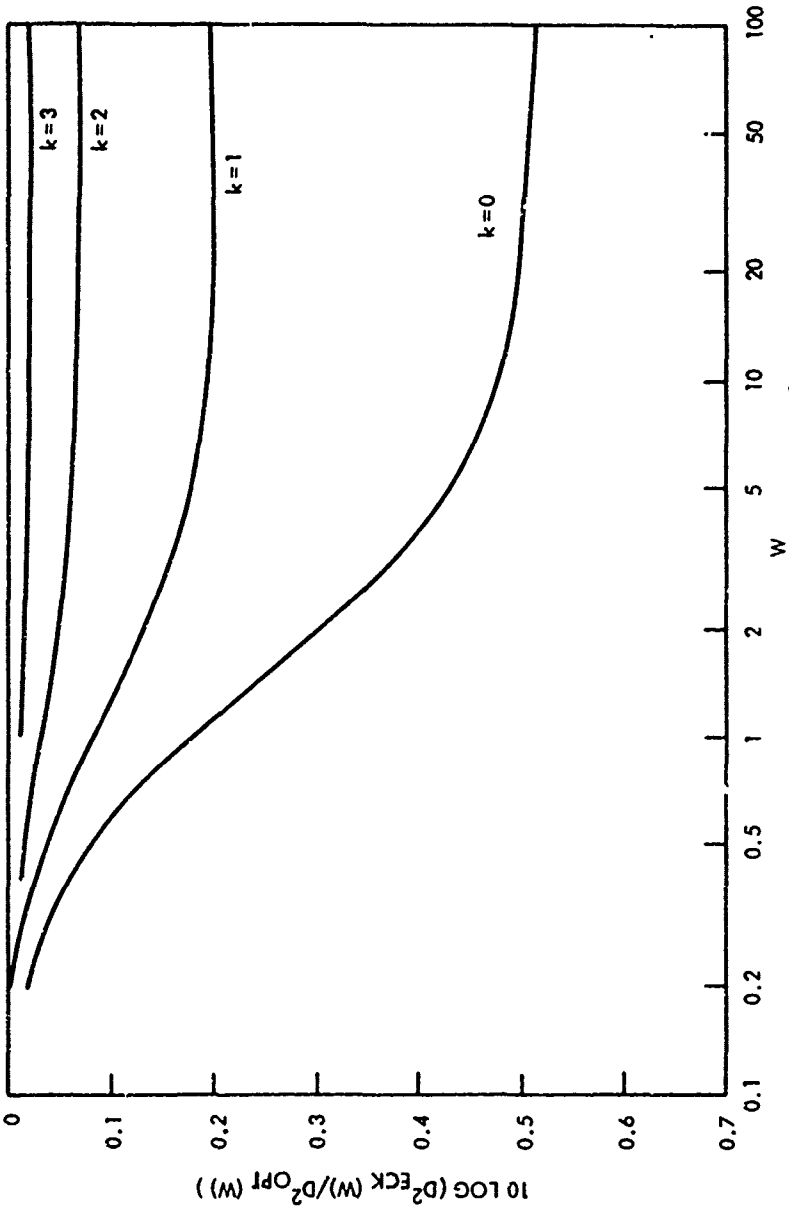

FIG. 5 SYSTEM DEGRADATION USING $|F_{ECK}|^2$, d = 2.

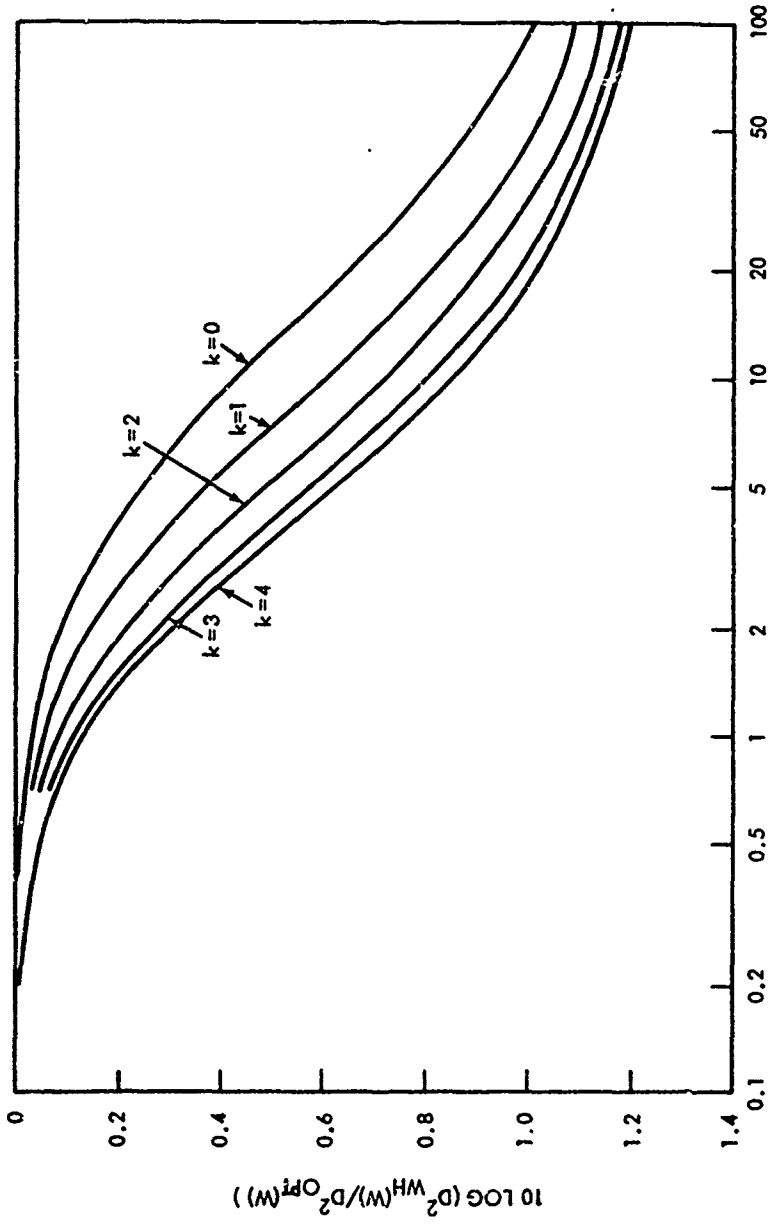

FIG. 6 SYSTEM DEGRADATION USING $|F_{WH}|^2$, d=1.

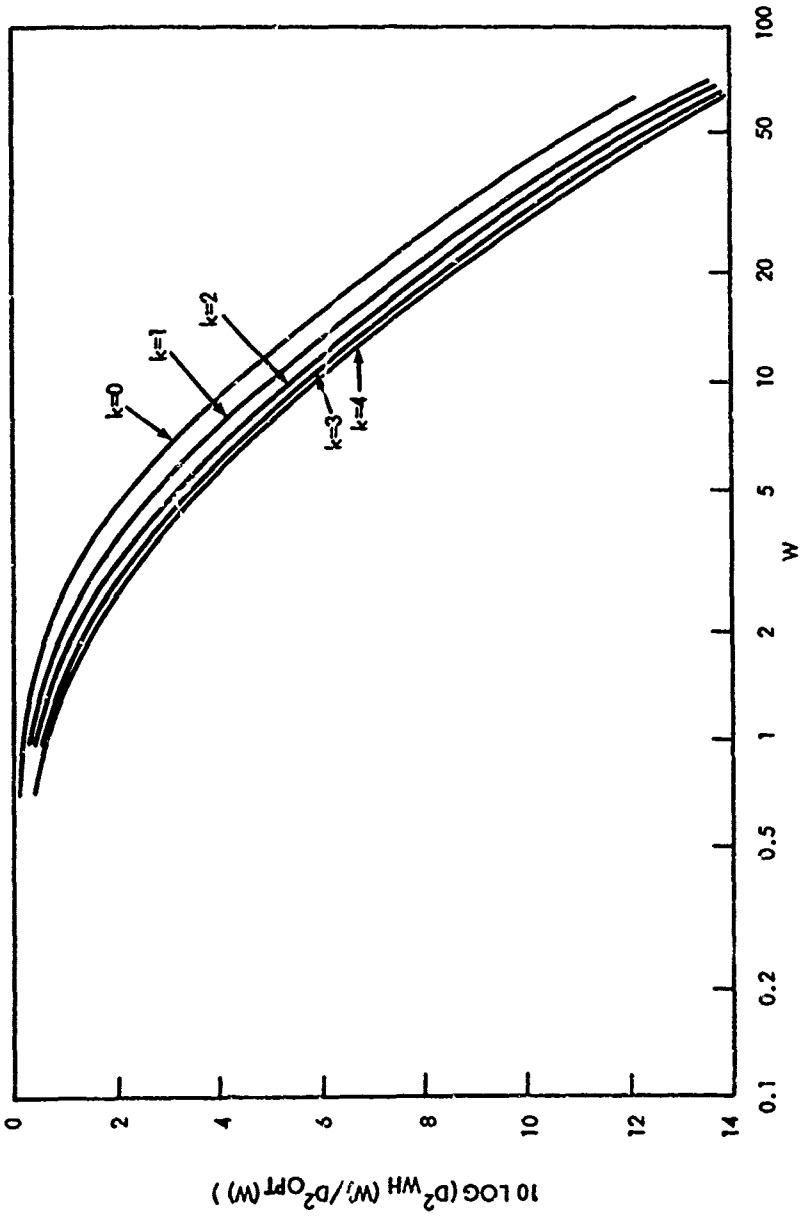

FIG. 7 SYSTEM DEGRADATION USING $|\rho_{WH}|^2$, d=2.

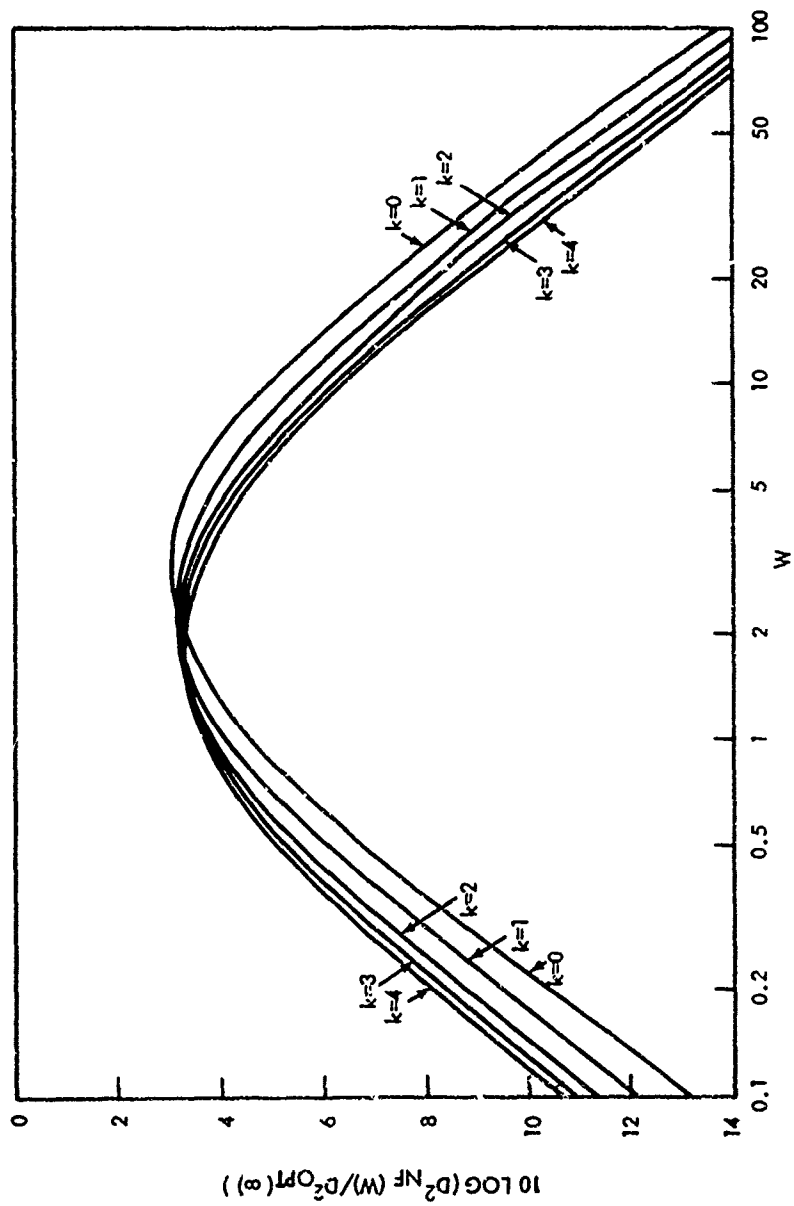

FIG. 8 SYSTEM DEGRADATION USING $|F_{NF}|^2$, d=2, a=2, b=0.

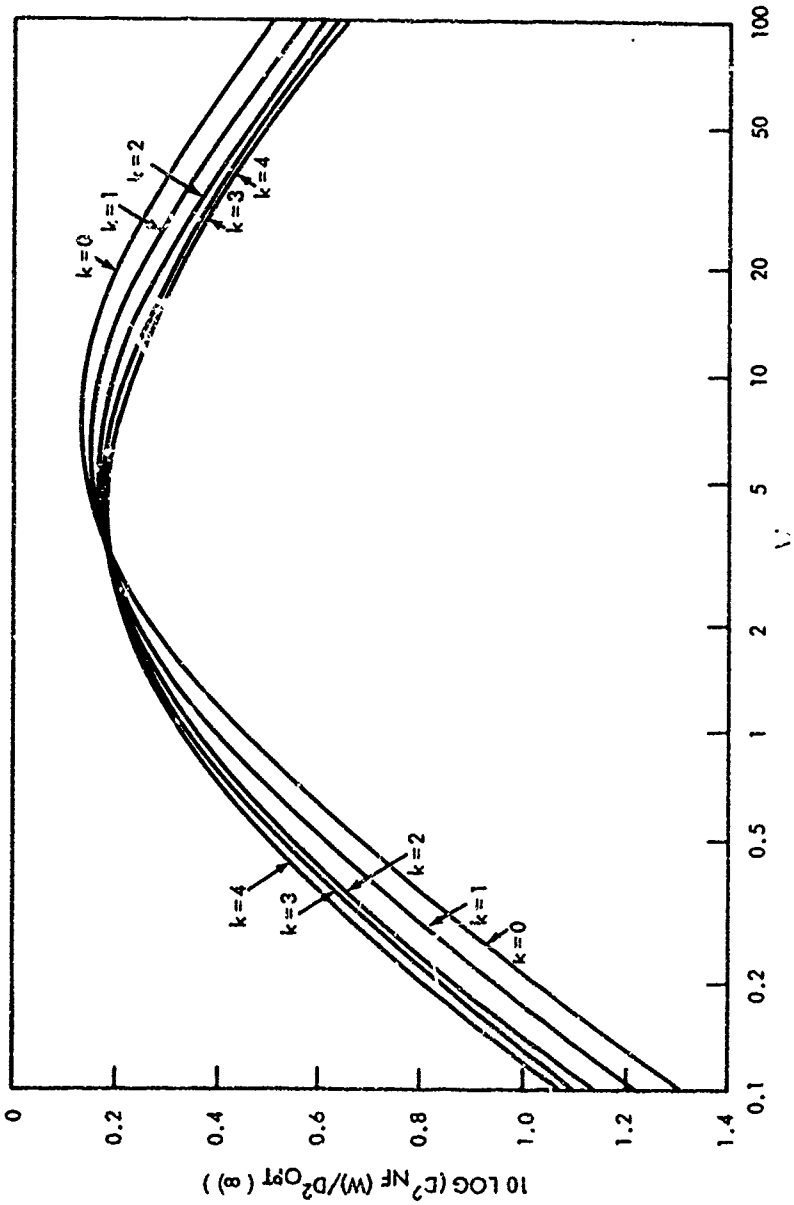

FIG. 9 SYSTEM DEGRADATION USING $|F_{NF}|^2$, d=2, a=3, b=1.

APPENDIX A

THE CRAMÉR-RAO MATRIX BOUND

The Cramér-Rao matrix bound (CRMB) is the generalization to the vector estimation case of the scalar Cramér-Rao bound. The CRMB is useful for the same reason that the scalar bound is useful--it is a bound for all estimators, and can be calculated from a knowledge of only the probability density function without specifying an estimator. It is easily derived, as follows:

If u, v, and w are vectors, and H is a matrix, and if only v and w are random, consider the quadratic form

$$0 \leq u^T(v-Hw)(v-Hw)^T u$$

$$= u^T(vv^T+Hww^TH^T-2Hwv^T)u. \qquad (A.1)$$

Assume $< ww^T >^{-1}$ exists, and let

$$H = < vv^T > < ww^T >^{-1}, \qquad (A.2)$$

Take the expectation of both sides of (A.1) and use (A.2). The result is the quadratic form

$$0 \leq u^T (< vv^T > - < vw^T >< ww^T >^{-1} < wv^T >)u. \qquad (A.3)$$

(A.3) shows that the matrix in the quadratic form of (A.3) is nonnegative definite, since u is an arbitrary vector. Thus

$$< vv^T > \geq < vw^T > < ww^T >^{-1} < wv^T > . \qquad (A.4)$$

If f is a scalar function of the vector y, let the gradient of f with respect to y be written as a row vector

$$\text{grad } f = \left(\frac{\partial f}{\partial y_1} , \frac{\partial f}{\partial y_2} , \ldots, \frac{\partial f}{\partial y_n}\right). \qquad (A.5)$$

If g is a column vector, let grad g be that matrix in which the i-th row is the gradient (with respect to y) of the i-th component of g.

$$\text{grad } g = \begin{bmatrix} \text{grad } g_1 \\ \vdots \\ \text{grad } g_n \end{bmatrix} . \qquad (A.6)$$

Consider a nonrandom vector estimation problem. The parameter vector to be estimated, y, and the vector of observations, x, are the arguments of the probability density function $p(x|y)$. Let $\tilde{y}(= \tilde{y}(x))$ be an estimate for y based on the observation vector x. With the gradients taken with respect to the components of y,

$$\text{grad} < \tilde{y} > = \text{grad} \int \tilde{y}\, p(x|y)\, dx$$

$$= \int \tilde{y}\, \text{grad}\, p(x|y)\, dx$$

$$= \int \tilde{y} (\text{grad ln}\, p(x|y))\, p(x|y)\, dx$$

$$= < \tilde{y}\, \text{grad ln}\, p(x|y) > . \qquad (A.7)$$

Also,

$$< y\, \text{grad ln}\, p(x|y) > = \int y\, \text{grad}\, p(x|y)\, dx$$

$$= y\, \text{grad} \int p(x|y)\, dx$$

$$= \underline{0} \ . \qquad (A.8)$$

From (A.7) and (A.8),

$$\text{grad} < \tilde{y} > = < (\tilde{y}-y)\, \text{grad ln}\, p(x|y) > . \qquad (A.9)$$

If the estimate $\tilde{y}$ has the bias b, that is, $< \tilde{y} > = y + b$, then

$$\text{grad} < \tilde{y} > = I + \text{grad}\, b \ . \qquad (A.10)$$

In (A.4) let $v = (\tilde{y}-y)$ and $w^T = \text{grad ln}\, p(x|y)$. Then $< vw^T > = I + \text{grad}\, b$, in view of (A.10), and (A.4) becomes

$$< (\tilde{y}-y)(\tilde{y}-y)^T > \geq (I + \text{grad}\, b)$$

$$\cdot < (\text{grad ln}\, p(x|y))^T (\text{grad ln}\, p(x|y)) >^{-1}$$

$$\cdot (I + \text{grad}\, b)^T \ . \qquad (A.11)$$

This is known as the Cramér-Rao Matrix Bound (CRMB). The matrix

$$(\text{FIM}) = \langle (\text{grad } \ln p(x|y))^T (\text{grad } \ln p(x|y)) \rangle \tag{A.12}$$

is called the Fisher Information Matrix (FIM).

A relevant question is whether it is possible to achieve equality in (A.11). This is possible only if there is equality in (A.1), which occurs only if $v = Hw$. Note that H as defined in (A.2) is not random, that is, it is not a matrix function of the data. H is at most a matrix function of y. Assume H^{-1} exists. Then equality in (A.11) is possible if and only if

$$\text{grad } \ln p(x|y) = H^{-1}(y)(\tilde{y}(x)-y) \ . \tag{A.13}$$

That is, grad ln p(x|y) can be factored into a matrix product, where $H^{-1}(y)$ is a matrix that does not depend on the data.

The development above is patterned after Balakrishnan [12].

An alternate form of the FIM is often more convenient to use than the right side of (A.13). It is derived as follows:

$$\int p(x|y)\, dx = 0 \,. \tag{A.14}$$

$$\begin{aligned} \text{grad} \int p(x|y)\, dx &= \int \text{grad}\, p(x|y)\, dx \\ &= \int (\text{grad ln}\, p(x|y))\, p(x|y)\, dx \\ &= \underline{0} \,. \end{aligned} \tag{A.15}$$

$$\begin{aligned} &\text{grad} \int (\text{grad ln}\, p(x|y))^T\, p(x|y)\, dx \\ &= \int \text{grad}[(\text{grad ln}\, p(x|y))^T\, p(x|y)]\, dx \\ &= \int [\text{grad}(\text{grad ln}\, p(x|y))^T]\, p(x|y)\, dx \\ &+ \int (\text{grad ln}\, p(x|y))^T (\text{grad ln}\, p(x|y))\, p(x|y)\, dx \\ &= \underline{0} \,. \end{aligned} \tag{A.16}$$

That grad $p(x|y)$ = (grad ln $p(x|y)$) $p(x|y)$ was used in (A.16). From the last equality in (A.16) it follows that

$$\begin{aligned} (\text{FIM}) &= -\int [\text{grad}(\text{grad ln}\, p(x|y))^T]\, p(x|y)\, dx \\ &= - < \text{grad}(\text{grad ln}\, p(x|y))^T > \,, \end{aligned} \tag{A.17}$$

which is the alternate form of the FIM.

APPENDIX B

COMPLEX GAUSSIAN RANDOM VARIABLES

Let x and y be real Gaussian random n-vectors, with means $\bar{x}$ and $\bar{y}$, respectively. The joint probability density function for x and y is

$$p(x,y) = (2\pi)^{-n} (\det P)^{-1/2} \exp\left\{-\frac{1}{2}\begin{bmatrix} x-\bar{x} \\ y-\bar{y} \end{bmatrix}^T P^{-1} \begin{bmatrix} x-\bar{x} \\ y-\bar{y} \end{bmatrix}\right\}, \tag{B.1}$$

where

$$P = \left\langle \begin{bmatrix} x-\bar{x} \\ y-\bar{y} \end{bmatrix} \begin{bmatrix} x-\bar{x} \\ y-\bar{y} \end{bmatrix}^T \right\rangle = \begin{bmatrix} V_{xx} & V_{xy} \\ V_{yx} & V_{yy} \end{bmatrix} . \tag{B.2}$$

Let tne complex random vector $z \equiv x + jy$, where the real and imaginary parts are distributed according to (B.1). Under certain practical and important conditions the joint density function can be written as a function of the complex random vector z, rather than as a function of both x and y. This is desirable in that expectations of functions of z, and also investigations of the properties of the likelihood function, become much more tractable.

It is convenient to state certain matrix results. Consider matrices of the form $R = \begin{bmatrix} V & -W \\ W & V \end{bmatrix}$ and $C = V + jW$, where V and W are real square n x n matrices. Then the following statements are true:

1. Matrices of the form R and C are isomorphic under matrix addition and matrix multiplication.

2. R is symmetric $\Leftrightarrow$ C is Hermitian ($C = C^{T*}$)

$$\Leftrightarrow V = V^T \text{ and } W = -W^T.$$

3. $C^{-1} = A + jB$ exists $\Leftrightarrow R^{-1} = \begin{matrix} A & -B \\ B & A \end{matrix}$ exists.

4. R is orthogonal ($R^T = R^{-1}$) $\Leftrightarrow$ C is unitary ($C^{-1} = C^{*T}$).

5. R is symmetric positive definite $\Leftrightarrow$ C is Hermitian positive definite.

6. If $V = V^T$ and $W = -W^T$, if x and y are n vectors, and if $z = x + jy$, then

$$(x^T \; y^T) \begin{bmatrix} V & -W \\ W & V \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = z^T C z^* \; , \tag{B.3}$$

and

$$(x^T \; y^T) \begin{bmatrix} V & -W \\ W & V \end{bmatrix}^{-1} \begin{bmatrix} x \\ y \end{bmatrix} = z^T C^{-1} z^* . \tag{B.4}$$

7. If $V = V^T$ and $W = -W^T$, then

$$\det R = (\det C)^2 \; . \tag{B.5}$$

Statements 1 through 6 are easy to verify. Only statement 7 requires some justification. In Mathews and Walker [13, p. 155] it is shown that a complex matrix C can be diagonalized by a unitary matrix T if and only if C and C^{*T} commute ($C^{*T}C = CC^{*T}$). If C is Hermitian ($C = C^{*T}$) then C and C^{*T} obviously commute. Let T be the unitary transformation that diagonalizes the Hermitian matrix C. Then $T^{-1}CT = D$, where D is a diagonal matrix. Since det T =

$(\det T^{-1})^{-1}$, $\det C = \det D$. Let $\tilde{T}$, $\tilde{C}$, and $\tilde{D}$ be the matrices isomorphic to T, C, and D, respectively. Then $\tilde{D} = (\tilde{T})^{-1}\tilde{C}\tilde{T}$, so that $\det \tilde{D} = \det \tilde{C}$. But $\tilde{C} = R$, and $\tilde{D} = \begin{bmatrix} D & O \\ O & D \end{bmatrix}$, so that $\det \tilde{D} = (\det D)^2$. Thus $\det R = (\det C)^2$.

Thus, if $z = x + jy$, where x and y are Gaussian n-vectors distributed by (B.1), and if

$$P = \frac{1}{2}\begin{bmatrix} V & -W \\ W & V \end{bmatrix} , \qquad (B.5)$$

where $V = V^T$ and $W = -W^T$, then using (B.4) and (B.5) the density function (B.1) can be written as

$$p(z) = \pi^{-n} (\det C)^{-1} \exp\{-(z-\bar{z})^{T*} C^{-1} (z-\bar{z})\} \qquad (B.6)$$

where $C = V + jW$ and $\bar{z} = \bar{x} + j\bar{y}$. The complex vector argument of (B.6) is called a <u>complex Gaussian random vector</u>.

Complex Gaussian random vectors and their properties have been extensively studied by Goodman [4]. In [4] Goodman develops, among other things, results pertaining to characteristic functions, maximum likelihood estimation of C, and distribution functions for the maximum likelihood estimates.

Whether or not the complex random variables are distributed by (B.6), the relation

$$\langle z_1 z_2 z_3 z_4 \rangle = \langle z_1 z_2 \rangle \langle z_3 z_4 \rangle + \langle z_1 z_3 \rangle \langle z_2 z_4 \rangle + \langle z_1 z_4 \rangle \langle z_2 z_3 \rangle - 2 \langle z_1 \rangle \langle z_2 \rangle \langle z_3 \rangle \langle z_4 \rangle \qquad (B.7)$$

holds if the real and imaginary parts of the complex random variables in (B.7) are Gaussian random variables.

REFERENCES

[1] M. J. Levin, "Power Spectrum Parameter Estimation," IEE Trans. Inform. Theory, Vol. IT-11, pp. 100-107, Jan 1965.

[2] R. O. Harger, "Statistically Optimum Optical Data Processing with Automatic Focus Estimation," IEEE Trans. Aerospace Electronic Systems, Vol. AES-2, No. 6, Nov 1966, pp. 665-675.

[3] V. H. MacDonald and P. M. Schultheiss, "Optimum Passive Bearing Estimation in a Spatially Incoherent Noise Environment," J. Acou. Soc. Am., Vol. 46, No. 1 (Part 1), pp. 37-43.

[4] N. R. Goodman, "Statistical Analysis Based on a Certain Multivariate Complex Distribution (An Introduction)," Ann. Math. Stat., 34, pp. 152-177.

[5] P. M. Woodward, Probability and Information Theory, With Applications to Radar, Pergamon Press, London, 1953.

[6] A. C. Bryson and Y. C. Ho, Applied Optimal Control, Blaisdell Publishing Co., Waltham, Massachusetts, 1969.

[7] R. Deutsch, Estimation Theory, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1965.

[8] G. M. Jenkins and D. G. Watts, Spectral Analysis and its Applications, Holden-Day, Inc., San Francisco, 1968

[9] D. G. Luenberger, Optimization By Vector Space Methods, John Wiley and Sons, Inc., New York, 1969.

[10] C. Eckart, University of California, Marine Physical Laboratory of Scripps Institute of Oceanography, SIO Ref. 52-11 (1952).

[11] R. J. Whelchel, "Optimization of Conventional Passive Sonar Detection Systems," Naval Avionics Facility, Indianapolis, Ind., NAFI TR-1440 (June 1969).

[12] A. V. Balakrishnan, "Filtering and Prediction Theory," in Communication Theory, McGraw-Hill Book Co., New York, 1968.

[13] J. Mathews and R. L. Walker, Mathematical Methods of Physics, W. A. Benjamin, Inc., New York, 1964.